DANIELE DE FREITAS PARMA

ANATOMICAL, MORPHOPHYSIOLOGICAL AND MOLECULAR CHARACTERIZATION OF BRAZILIAN SPECIES OF CLEOMACEAE

Thesis submitted to the Botany Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Advisor: Adriano Nunes Nesi

Co-advisor: Marcelo Gomes Marçal Vieira Vaz

VIÇOSA - MINAS GERAIS
2021

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

Parma, Daniele de Freitas, 1993-
P253a Anatomical, morphophysiological and molecular characterization of
2021 Brazilian species of Cleomaceae / Daniele de Freitas Parma. - Viçosa, MG, 2021.
1 tese eletrônica (204 f.): il. (algumas color.).

Orientador: Adriano Nunes Nesi.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Biologia Vegetal, 2021.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2021.209
Modo de acesso: World Wide Web.

1. *Cleoserrata*. 2. *Gynandropsis*. 3. *Tarenaya*. I. Nesi, Adriano Nunes, 1975- . II. Universidade Federal de Viçosa. Departamento de Biologia Vegetal. Programa de Pós-Graduação em Botânica. III. Título.

CDD 22. ed. 583

Bibliotecário(a) responsável: Renata de Fátima Alves CRB6/2578

DANIELE DE FREITAS PARMA

ANATOMICAL, MORPHOPHYSIOLOGICAL AND MOLECULAR CHARACTERIZATION OF BRAZILIAN SPECIES OF CLEOMACEAE

Thesis submitted to the Botany Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: August 27, 2021.

Assent:

Daniele de Freitas Parma

Daniele de Freitas Parma
Author

Adriano Nunes Nesi

Adriano Nunes Nesi
Advisor

ACKNOWLEDGMENTS

I would like to thank the Universidade Federal de Viçosa (UFV), mainly to the Botany Graduate Program, for the support and for providing me all conditions required during the development of my research. I am also grateful to the scholarships provided by CAPES (Coordination for the Improved of Higher Level Personnel) and FAPEMIG (Foundation for Research Assistance of the Minas Gerais State) and CNPq (National Council for Scientific and Technological Development) for financing the project.

I sincerely thank my adviser prof. Adriano Nunes Nesi for his guidance, support, trust and advice. I am also grateful to Dr. Marcelo Vaz for the friendship and for helping me with discussions and advices for this thesis. I would like to thank Prof. Wagner Araújo, Prof. Andreas Weber and Prof. Erik Schranz for helping me with discussions methodological and the results of this thesis. I want to thank the members of my evaluation committee Prof. Samuel Martins, Prof. Andreas Weber, Prof. Wagner Araújo, Prof. Jéferson Fregonezi and Dr. Marcelo Vaz.

I would also like to thank all colleagues at UCP, especially those who have followed the development of the work more closely, such as Dora, David, Jaciara, Regina, Rina, Italo, Samuel, Leo, Priscilla, Kaik and Sandy.

I also want to thank my great friends who mobilized and helped me to get seeds to develop the work. As well as those who accompanied me in fieldwork. Thank you all very much. I would also like to thank IEF for the license granted, as well as the herbaria that received me and lent us/donated Cleomaceae materials.

My special thanks also to the friends of botany, Ronaldo, Valentin, Genilson, Lívia, Evandro and Celso for their friendship and discussions about the work. I also thank the staff of the genetics laboratory, especially Prof. Well and Jessica who helped me with the analysis.

I also express here my gratitude to all the professors, technicians, and colleagues at the Botany Graduate Program for the support. Many thanks for all other people who in some way had ever helped or supported me during my doctoral studies and whose names are not mentioned here.

I sincerely would like to express my gratitude to my family. My parents, Maury and Nilva and my sister Driele. My husband, Luiz, for all the companionship, love and discussions about science. My dear son, who despite being a little baby, teaches me every day about love, patience and perseverance.

ABSTRACT

PARMA, Daniele de Freitas, D.Sc., Universidade Federal de Viçosa, August, 2021. **Anatomical, Morphophysiological and Molecular Characterization of Brazilian Species of Cleomaceae.** Advisor: Adriano Nunes Nesi. Co-advisor: Marcelo Gomes Marçal Vieira Vaz.

Cleomaceae is one of the 17 families of the order Brassicales, which comprises 26 genera and about 270 species. This cosmopolitan family has a great morphological and physiological diversity and due it is a sister group of Brassicaceae, to which the model plant *Arabidopsis* is inserted, it has been the subject of studies in several areas. However, despite being a promising family, there are few species characterized, so far only 15 have been physiologically characterized. Thus, this work aimed to characterize species of Cleomaceae occurring in Brazil. In this sense, visits to herbariums and collections were carried out in various regions in order to better understand the identity, ecology and distribution of each species. Thus, 22 species were sampled. Morphological and molecular characterization, based on five markers, showed that the *Tarenaya* series are monophyletic (Chapter 1) and that *T. siliculifera* is morphologically distinct from the species of the genus, so the recognition of a new monotype genus has been proposed (Chapter 2). In addition, we characterized the 22 species in terms of genome size, molecular phylogeny, anatomical, biochemical, and physiological characteristics (Chapter 3). In addition to determining the photosynthetic parameters, biochemical analyzes were performed on the leaves throughout the day (Chapter 4). Interestingly, the C_3 species groups exhibit clear variation in anatomical, physiological and biochemical parameters, suggesting a large natural variation among these species. Furthermore, unrecognized variation in leaf structure and physiology within the species studied indicates varying degrees of development of the C_3-C_4 photosynthetic mechanism. Furthermore, the Brazilian Cleomaceae accessions characterized here exhibit a high genetic diversity and may represent different stages of development of the photosynthetic mechanism from C_3 to C_4. Regarding the characterization of the sexual type of each flower (Chapter 5), the 20 morphospecies formed three four groups of sexual expression: andromonoicia, hermaphrodites, and polygamous. In addition, dichogamy and distyly were also observed, however, the species studied are self-compatible. Finally, a compilation of all available information about Cleomaceae is carried out in Chapter 6 in order to identify the characteristics of each group and indicate the advances in each area with the family.

Keywords: *Cleoserrata*. *Gynandropsis*. Molecular phylogeny. Photosynthesis. *Tarenaya*.

RESUMO

PARMA, Daniele de Freitas, D.Sc., Universidade Federal de Viçosa, agosto de 2021. **Caracterização Anatômica, Morfisiológica e Molecular de Espécies Brasileiras de Cleomaceae**. Orientador: Adriano Nunes Nesi. Coorientador: Marcelo Gomes Marçal Vieira Vaz.

Cleomaceae é uma das 17 famílias de Brassicales, compreende 26 gêneros e 270 espécies. Esta família cosmopolita possui uma grande diversidade morfológica e fisiológica e por ser grupo-irmão de Brassicaceae, ao qual a planta modelo *Arabidopsis* está inserida, tem sido alvo de estudos em diversas áreas. No entanto, apesar de ser uma família promissora, são poucas espécies caracterizadas, até o momento somente 15 foram caracterizadas fisiologicamente. Sendo assim, este trabalho objetivou caracterizar espécies de Cleomaceae ocorrentes no Brasil. Nesse sentido, visitas a herbários e coletas foram realizadas em várias regiões do país em busca de compreender melhor a identidade, ecologia e distribuição de cada espécie. Dessa forma, 22 morfoespécies foram amostradas. A caracterização morfológica e molecular, baseada em cinco marcadores (ITS, matK, ndhF, ycF, rps), mostrou que as series de *Tarenaya* são monofileticas (Capítulo 1) e que *T. siliculifera* é morfologicamente distinta das espécies do gênero, sendo então proposto o reconhecimento de um novo gênero monotipo (Capítulo 2). Além disso, caracterizamos as 22 morfoespécies em termos de tamanho do genoma, filogenia molecular, características anatômicas, bioquímicas e fisiológicas (Capítulo 3). Nesse sentido, além da determinação dos parâmetros fotossintéticos, foram realizadas análises bioquímicas nas folhas ao longo do dia (Capítulo 4). Curiosamente, os grupos de espécies C_3 exibem clara variação nos parâmetros anatômicos, fisiológicos e bioquímicos, sugerindo uma grande variação natural entre essas espécies. Além disso, a variação não reconhecida na estrutura e fisiologia da folha dentro das espécies estudadas indica vários graus de desenvolvimento do mecanismo fotossintético C_3-C_4. Em relação a caracterização do tipo sexual de cada flor (Capítulo 5), as 20 morfoespécies analizadas formaram três grupos de expressão sexual: andromonoicia, hermafroditas e polígamas. Além disso, também foram observadas dicogamia e distilia (hercogamia), no entanto, as espécies estudadas são autocompatíveis. Finalmente, é realizado um compilado das informações disponíveis sobre Cleomaceae no Capítulo 6 de forma a identificar as características de cada grupo e identificar os avanços em cada área com a família.

Palavras-chave: *Cleoserrata*. *Gynandropsis.* Filogenia molecular. Fotossíntese. *Tarenaya.*

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

SUMMARY

GENERAL INTRODUCTION

CHAPTER 1

CHAPTER 2

CHAPTER 4

CHAPTER 6

GENERAL INTRODUCTION

1.1 Taxonomy and Systematics of Cleomaceae

Cleomaceae is one of the 17 botanical families belonging to the order Brassicales (Stevens 2001). It is a relatively small family comprising 26 genera (Soares Neto et al., 2020) and about 270 species (Bayat et al., 2018). Cleomaceae has been the subject of several taxonomy and phylogeny studies, even when it was considered as a subfamily of Capparaceae (Cleomoideae) (Iltis, 1957; Hall et al., 2002; Sanchez-Acebo, 2005; Inda et al., 2008; Feodorova et al., 2010). It is important to mention that some studies questioned the position of Cleomoideae as well as the circumscription of the genus *Cleome* s.l. (Iltis, 1957; Hall et al., 2002; Sanchez-Acebo, 2005; Inda et al., 2008; Feodorova et al., 2010). Accordingly, after the work of Iltis et al., (2011) recognizing Cleomaceae as an independent family, phylogenetically close to Brassicaceae and Capparaceae (sister families), many studies have been performed with the purpose of delimiting *Cleome* s.s. and other genera, making them monophyletic entities (Roalson et al., 2015; Barrett et al., 2017; Thulin and Roalson, 2017; Soares-Neto et al., 2018).

Patchell et al., (2014) conducted a molecular phylogeny study using five markers (ITS, matk, ycf, ndhF, rps3) from three genomes (nuclear, chloroplast and mitochondrial), aiming to understand the cluster *Cleome* s.s.. Based on these data, they were able to separate Cleomaceae into 15 clusters. Since then, many studies of taxonomy and systematics of have used this work to validate previously proposed taxonomic changes based only on morphological characteristics (*e.g.* Soares-Neto et al., 2018). In this sense, following these works, some clusters described by are now considered as monophyletic groups (Patchell et al., (2014). In addition to molecular traits, the main morphological characteristics used to separate the species in Cleomaceae were: growth form, pubescence, leaflet shape, petal size and stamen number (Roalson et al., 2015; Barrett et al., 2017; Thulin and Roalson, 2017; Soares-Neto et al., 2018).

The members of Cleomaceae are separated into Neocleome (New World) and Eucleome (Old World) (Stevens 2001). Most recent taxonomic studies have been performed with the Eucleome group (Iltis and Cochrane, 2014; Roalson et al., 2015; Barrett et al., 2017; Thulin and Roalson, 2017). However, Neocleome has been studied for a long time, mainly by Iltis (1952; 1957). Thus, according to Iltis (1957) *Cleome* is more abundant in tropical and subtropical New World regions, with two main centers of distribution in Mexico and South America that apparently included the most 'primitive' taxa. According to the same author, the basal divergences in the genus and relationships between Old World and New World have not yet been clarified, as can also be seen in a more recent study (Feodorova et al., 2010). The works of Iltis were extremely important for the understanding/knowledge, especially, of the New World Cleomaceae. Only with morphological and ecological data he recognized groups that were later confirmed by molecular data. In this sense, especially after the work of Patchell et al., (2014) some sections proposed by Iltis (1952) have recently been elevated to genera, such as *Melidiscus, Physostemon, Tarenaya and Cleoserrata*. It is noteworthy that these changes had a molecular basis (Patchell et al., 2014), although the studies proposing such taxonomic changes in the family did so based only on morphology.

The Brazilian territory is considered as a diversity center of Cleomaceae, with a record of 34 species (Flora of Brazil 2021). *Tarenaya*, with 22 species, is the most diverse Cleomaceae genus in the country, followed by *Physostemon* (4 spp.), *Dactylaena* (3 spp.), *Cleoserrata* (2 spp.), *Haptocarpum* (1 spp.), *Gynandropsis* (1 spp.) and *Melidiscus* (1 spp.) (Flora of Brazil 2021). The genus *Tarenaya*, besides being the most specious, has the largest taxonomic conflicts. For example, the complex "*Tarenaya spinosa*", that, according to Iltis (1952), is composed of 10 species (see Soares-neto et al., 2019). Thus, "*Tarenaya spinosa*" is difficult to identify due to the high polymorphy of the species (Iltis 1952). But, recently Soares-Neto et al., (2018, 2019) solved this dilema delimiting the geographic distribution, as well as the main diagnostic characteristics for the recognition of each species.

It should be noted that *Tarenaya* comprises four groupings, Spinosa, Hemiscola, Parviflora and Rosea, and until recently, the genus was considered polyphyletic (Patchell et al., 2014). According to Feodorova et al., (2010), the polyphilia of *Tarenaya* should be solved by including *Hemiscola* within *Tarenaya* or by breaking this genus into more genera (possibly according to the series of *Tarenaya*). However, morphological characters associated with these clusters need to be explored in more detail in order to determine the best approach (Feodorova et al., 2010). Thus, Soares-Neto et al., (2018) described 25 new combinations for the genus *Tarenaya*. As consequence, species that were previously recognized as *Cleome* spp. or *Hemiscola* spp. were transferred to the genus *Tarenaya*. It is noteworthy that this study is only based on morphology, considering mainly the synapomorphies of *Tarenaya* (spines and prickles of epidermal origin, seeds with an aril and a large cleft). After taxonomic modifications performed by Soares-Neto et al., (2018), *Tarenaya* became a monophyletic group (Soares-Neto et al., 2020).

Tarenaya is a very morphologically diverse group and comprises 37 species (Soares-Neto et al., 2020), being one of the most specious genera of the family. Thus, in Chapter 1, we analyze all the native species of this genus, either by exsiccates or botanical collections, enabling us to described the main characteristics to separate them in their respective series. In addition, we analyzed some species in a molecular context, applying phylogenetic analysis based on five molecular markers. These specimens were also evaluated and classified for conservation status according to criteria adopted by the IUCN. Based on the results of Chapter 1, we propose the recognition of a novel genus into Cleomaceae. Unlike Soares-Neto et al., (2018), we believe, based on molecular and morphological data, that *Tarenaya* should be separated into small clusters. Although most of the series of *Tarenaya* share the synapomorphies of the genus (except *T. siliculifera*), morphological differences such as floral morphology, pollination strategies, seed characters and putative chromosome numbers (Iltis & Cochrane, 2007) stand out in comparison to the similarities. In addition, molecular data support these taxonomic modifications, leading each series to be recognized as a monophyletic cluster. Therefore, in Chapter 2 we reinforce these proposed modifications for *Tarenaya*, by elevating the species *T. siliculifera* to a new genus – *Neotarenaya siliculifera*.

1.2 Evolution of the C_4 Photosynthetic Mechanism

The world demand for food has increased exponentially over the years, mainly with population growth (Southgate, 2009). Despite of agriculture has gone through a revolution and today, many technologies at its disposal, the arable areas have decreased and, the world has been suffering with climate change that alter the local precipitation and temperatures (United Nations 2017; Valin et al., 2014). In that regard, it has become apparent that photosynthetic efficiency and capacity must be increased *per* unit leaf area to improve yield potential (Long et al., 2015). Thus, there is a need to accelerate the understanding of the photosynthetic process to allow informed and guided improvements via genetic engineering (Long et al., 2015). In this way, there are studies in several photosynthesis related processes, such as light capture, light energy conversion, carbon capture and conversion, smart canopy (see review Evans, 2013; Ort et al., 2015), as well as leaf anatomical traits (Ren et al., 2019), aiming at increasing photosynthetic rates. However, what has been shown to be more interesting is the C_4 photosynthetic pathway which is characterized by high photosynthetic rate and high nitrogen and water use efficiencies compared to plants with the C_3 photosynthetic mechanism (Mitchell and Sheehy, 2006). Accordingly, improvements of photosynthesis in C_3 crops include its conversion to C_4 metabolism that could increase the light conversion efficiency by 5–60% compared with the best cultivars (Long et al., 2015).

The C_4 mechanism has Kranz anatomy and requires the coordinated functioning of mesophyll and bundle-sheath cells in leaves and is characterized by a CO_2 concentrating mechanism which allows Rubisco, located in the bundle sheath cells (BSC), to operate at high CO_2 partial pressures, reducing photorespiration rates (Caemmerer, 2021). In the mesophyll, CO_2 is initially fixed by phosphoenolpyruvate (PEP) carboxylase into C_4 acids, which are then decarboxylated in the bundle sheath to supply CO_2 for Rubisco (Furbank, 2011). Both the structure of the bundle-sheath wall and the relative biochemical capacities of the C_3 cycle in the bundle sheath and C_4 acid cycle contribute to the high CO_2 partial pressure in the bundle sheath (Caemmerer, 2021).

During the evolution of the C_4 mechanism from C_3 species, a combination of complex anatomical and biochemical specialization was required (Sage 2004; Gowik and Westhoff, 2011). This evolutionary process of transition from C_3 to C_4 photosynthetic mechanism occurred independently in more than 65 angiosperms taxa (Sage 2004). Moreover, the evolution of C_4 photosynthesis is not a directional process and occurred in a series of events, such as development of larger BSC, increased leaf venation, restriction of glycine decarboxylase to BSC and establishment of photorespiratory carboxylase cycle, as well as enhancement of PEP carboxylase activity, establishment of C_4 cycle and optimization of the C_4 syndrome (Gowik and Westhoff, 2011; Sage, 2021).

Although all genes important for the C_4 pathway are expressed at relatively low levels in C_3 leaves (Gowik and Westhoff, 2011; Sage, 2004), the mechanism for recruitment of these genes into the C_4 pathway remains to be fully elucidated (Bergh et al., 2014). Noteworthy, the majority of studies on evolution of the C_4 mechanism has been conducted with plants phylogenetically distant to the model *Arabidopsis thaliana* (*e.g. Flaveria* – Asteraceae) (Gowik et al., 2011; Schulze et al., 2013). In this sense, Cleomaceae, a family closely related to Brassicaceae, which has species with C_3, C_4 and intermediate C_3-C_4 photosynthetic mechanisms, is considered promising in studies aiming to understand the evolution of C_4 photosynthetic mechanism (Marshall et al., 2007; Voznesenskaya et al., 2007; Bayat et al., 2018; Reeves et al., 2018). Despite the greatest species diversity within the family, few Cleomaceae species have been sampled for physiological studies to date. To this end, the Chapters 3 and 4 aimed at characterizing and identifying the type of photosynthetic metabolism displayed by Brazilian Cleomaceae species and, thus, gaining insights into the evolution of C_4 photosynthesis in this family.

1.3 Floral Diversity

Floral symmetry and mating systems are an ecologically and evolutionarily significant feature of angiosperms (Donoghue et al., 1989; Endress, 2001; Sargent, 2004; Kissling and Barrett, 2013). Therefore, the monosymmetry is considered a key innovation, with important implications for understanding angiosperm diversification (Sargent, 2004) and has evolved multiple times from polysymmetric flowers (Donoghue et al., 1989; Endress, 2001). Additionally, the herkogamy, the spatial separation of sexual organs within flowers, is a widespread floral mechanism that is thought to be an adaptive trait reducing self-pollination and the deleterious consequences of inbreeding depression on progeny fitness (Kissling and Barrett, 2013). Although there are various forms of herkogamy, three major types are commonly recognized: approach herkogamy (stigmas positioned above or protruding beyond the anthers); reverse herkogamy (stigmas positioned below or behind the anthers); and reciprocal herkogamy (stigmas and anthers placed in complementary positions in two or more floral morphs - dystily) (Opedal, 2018). The non-herkogamous condition homostyly commonly occurs as a derived condition in many heterostylous groups as a result of the evolutionary breakdown of distyly (Naiki, 2012).

The molecular basis of monosymmetry and herkogamy, for example, remains largely unexplored across diverse species, which can be primarily attributable to the limited genomic resources that are available for those plant species, which are generally non-model organisms or are not phylogenetically close to the model plant (*e.g. A. thaliana*). As the closest predominantly monosymmetric clade to Brassicaceae, Cleomaceae is ideal for developmental genetics studies addressing the origins of monosymmetry, since the family is characterized by having monosymmetric flowers. In this way, the flowers have a disymmetric ground plan (four sepals, four petals, usually six stamens, and bicarpellate gynoecium). The monosymmetric flowers are plesiomorphic and represent the dominant state in the family with at least one reversal to polysymmetry (Hall et al., 2002). The family also provides a valuable opportunity to investigate the evolutionary relationships among the three main herkogamous conditions. Moreover, species of Cleomaceae exhibit considerable floral diversity and associated features (such as nectar and pollen). This diversity provides an opportunity to assess to what extent variation in floral traits may be associated with evolutionary transitions among herkogamic conditions, an approach that has proven to be valuable in other groups (*e.g.* Gentianaceae - Kissling and Barrett, 2013).

Nonetheless, studies on the floral morphology and reproductive biology of the family are still scarce. In this way, the Chapter 5 contributes to a better understanding of the morphological diversity and reproductive biology of 20 Brazilian Cleomaceae species, which were sampled from different Brazilian biomes. The necessity or benefit of pollinators and cross-pollination was evaluated for producing fruits, fertile seeds and vigorous progeny. Additionally, in Chapter 6, it is possible to see that the morphological and physiological diversity observed in the family allows advances in several areas of science.

REFERENCES

Barrett, R. L., Roalson, E. H., Ottewell, K., Byrne, M., Govindwar, S. P., Yadav, S. R., et al. (2017). Resolving Generic Boundaries in Indian-Australasian Cleomaceae: Circumscription of *Areocleome* , *Arivela* , and *Corynandra* as Distinct Genera. *Syst. Bot.* 42, 694–708. doi:10.1600/036364417X696401.

Bayat, S., Schranz, M. E., Roalson, E. H., and Hall, J. C. (2018). Lessons from Cleomaceae, the Sister of Crucifers. *Trends Plant Sci.* 23, 808–821. doi:10.1016/j.tplants.2018.06.010.

Bergh, E. Van Den, Külahoglu, C., Bräutigam, A., Hibberd, J. M., Weber, A. P. M., Zhu, X., et al. (2014). Current Plant Biology Gene and genome duplications and the origin of C 4 photosynthesis : Birth of a trait in the Cleomaceae. *Curr. Plant Biol.* 1, 2–9. doi:10.1016/j.cpb.2014.08.001.

Caemmerer, S. von (2021). Updating the steady state model of C4 photosynthesis. *bioRxiv*, 2021.03.13.435281. doi:10.1101/2021.03.13.435281.

Donoghue, M. J., Doyle, J. A., Gauthier, J., Kluge, A. G., and Rowe, T. (1989). The importance of fossils in phylogeny reconstruction. *Annu. Rev. Ecol. Syst. Vol. 20*, 431–460. doi:10.1146/annurev.es.20.110189.002243.

Evans, J. R. (2013). Improving Photosynthesis. *Plant Physiol.* 162, 1780. doi:10.1104/PP.113.219006.

Feodorova, T. A., Voznesenskaya, E. V, Edwards, G. E., Roalson, E. H. (2010). Biogeographic Patterns of Diversification and the Origins of C4 in Cleome (Cleomaceae). 35, 811–826. doi:10.1600/036364410X539880.

Flora of Brazil - Cleomaceae Bercht. & J.Presl Available at: http://floradobrasil.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/BemVindoConsultaPublicaConsultar.do?invalidatePageControlCounter=1&idsFilhosAlgas=%5B2%5D&idsFilhosFungos=%5B1%2C10%2C11%5D&lingua=&grupo=5&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=Cleomaceae+Bercht.+%26+J.Presl&formaVida=null&substrato=null&ocorreBrasil=QUALQUER&ocorrencia=OCORRE

&endemismo=TODOS&origem=TODOS&regiao=QUALQUER&estado=QUALQUER&ilhaOceanica=32767&domFitogeograficos=QUALQUER&bacia=QUALQUER&vegetacao=TODOS&mostrarAte=SUBESP_VAR&opcoesBusca=TODOS_OS_NOMES&loginUsuario=Visitante&senhaUsuario=&contexto=consulta-publica [Accessed August 2, 2021].

Furbank, R. T. (2011). Evolution of the C 4 photosynthetic mechanism: are there really three C 4 acid decarboxylation types? *J. Exp. Bot.* 62, 3103–3108. doi:10.1093/jxb/err080.

Gowik, U., and Bra, A. (2011). Evolution of C 4 Photosynthesis in the Genus Flaveria : How Many and Which Genes Does It Take to Make C4 ? 23, 2087–2105. doi:10.1105/tpc.111.086264.

Gowik, U., and Westhoff, P. (2011). The Path from C_3 to C_4 Photosynthesis 1. 155, 56–63. doi:10.1104/pp.110.165308.

Hall, J. C., Sytsma, K. J., and Iltis, H. H. (2002). Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. *Am. J. Bot.* 89, 1826–1842. doi:10.3732/ajb.89.11.1826.

Iltis, H.H. 1952. A revision of the NewWorld species of Cleome. Dissertation, Washington University, St. Louis, Missouri, U.S.A.

Iltis, H. H. (1957). Studies in the Capparidaceae. III. Evolution and Phylogeny of the Western North American Cleomoideae. *Ann. Missouri Bot. Gard.* 44, 77. doi:10.2307/2394679.

Iltis, H. H., and Cochrane, T. S. (2014). Studies in the Cleomaceae VI : A New Genus and Sixteen New Combinations for the Flora Mesoamericana Studies in the Cleomaceae VI : A New Genus and Sixteen New Combinations for the Flora Mesoamericana. doi:10.3417/2013017.

Iltis, H. H., Hall, J. C., Cochrane, T. S., and Sytsma, K. J. (2011). Studies in the Cleomaceae I. On the Separate Recognition of Capparaceae, Cleomaceae, and Brassicaceae [1]. *Ann. Missouri Bot. Gard.* 98, 28–36. doi:10.3417/2007017.

Inda, L. A., Torrecilla, Æ. P., Catala, Æ. P., and Ruiz-zapata, T. (2008). Phylogeny of Cleome L . and its close relatives Podandrogyne Ducke and Polanisia Raf . (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. 111–126. doi:10.1007/s00606-008-0026-y.

Kissling, J., and Barrett, S. C. H. (2013). Variation and evolution of herkogamy in Exochaenium (Gentianaceae): Implications for the evolution of distyly. *Ann. Bot.* 112, 95–102. doi:10.1093/aob/mct097.

Long, S. P., Marshall-Colon, A., and Zhu, X.-G. (2015). Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell* 161, 56–66. doi:10.1016/J.CELL.2015.03.019.

Marshall, D. M., Muhaidat, R., Brown, N. J., Liu, Z., Stanley, S., Griffiths, H., et al. (2007). Cleome , a genus closely related to Arabidopsis , contains species spanning a developmental progression from C 3 to C 4 photosynthesis. 886–896. doi:10.1111/j.1365-313X.2007.03188.x.

Mitchell, P. L., and Sheehy, J. E. (2006). Supercharging rice photosynthesis to increase yield. *New Phytol.* 171, 688–693. doi:10.1111/J.1469-8137.2006.01855.X.

Naiki, A. (2012). Heterostyly and the possibility of its breakdown by polyploidization. *Plant Species Biol.* 27, 3–29. doi:10.1111/J.1442-1984.2011.00363.X.

Opedal, Ø. H. (2018). Herkogamy, a Principal Functional Trait of Plant Reproductive Biology. *https://doi.org/10.1086/700314* 179, 677–687. doi:10.1086/700314.

Ort, D. R., Merchant, S. S., Alric, J., Barkan, A., Blankenship, R. E., Bock, R., et al. (2015). Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci.* 112, 8529–8536. doi:10.1073/PNAS.1424031112.

Patchell, M. J., Roalson, E. H., and Hall, J. C. (2014). Resolved phylogeny of Cleomaceae based on all three genomes. 63, 315–328.

PK, E. (2001). Evolution of floral symmetry. *Curr. Opin. Plant Biol.* 4, 86–91. doi:10.1016/S1369-5266(00)00140-0.

Reeves, G., Singh, P., Rossberg, T. A., Sogbohossou, E. O. D., Schranz, M. E., and Hibberd, J. M. (2018). Natural variation within a species for traits underpinning C 4 photosynthesis. *Plant Physiol.* 177, 504–512. doi:10.1104/pp.18.00168.

Ren, T., Weraduwage, S. M., and Sharkey, T. D. (2019). Prospects for enhancing leaf photosynthetic capacity by manipulating mesophyll cell morphology. *J. Exp. Bot.* 70, 1153–1165. doi:10.1093/JXB/ERY448.

Roalson, E. H., Hall, J. C., Ii, J. P. R., and Cardinal-mcteague, W. M. (2015). A revision of generic boundaries and nomenclature in the North American cleomoid clade (Cleomaceae). 205, 129–144.

Sage, R. F. (2004). The evolution of C4 photosynthesis. *New Phytol.* 161, 341–370. doi:10.1111/j.1469-8137.2004.00974.x.

Sage, R. F., and Sage, R. F. (2004). The evolution of C 4 photosynthesis. 341–370. doi:10.1046/j.1469-8137.2004.00974.x.

Sanchez-Acebo, L. (2005). A Phylogenetic Study of the New World Cleome (Brassicaceae, Cleomoideae) on JSTOR. *Ann. Missouri Bot. Gard.* 92, 179–201. Available at: https://www.jstor.org/stable/3298513?seq=1#metadata_info_tab_contents [Accessed February 27, 2019].

Sargent, R. D. (2004). Floral symmetry affects speciation rates in angiosperms. *Proc. R. Soc. B Biol. Sci.* 271, 603. doi:10.1098/RSPB.2003.2644.

Schulze, S., Mallmann, J., Burscheidt, J., Koczor, M., Streubel, M., Bauwe, H., et al. (2013). Evolution of C4 Photosynthesis in the Genus Flaveria : Establishment of a Photorespiratory CO 2 Pump. 25, 2522–2535. doi:10.1105/tpc.113.114520.

Soares Neto, R.L., S., Thomas, W. W., Regina, M., Barbosa, D. V., Roalson, E. H., et al. (2018). New combinations and taxonomic notes for Tarenaya (Cleomaceae). 1–6. doi:10.1590/0102-33062017abb0417.

Soares Neto, R.L., Thomas,W.W., Barbosa, M.R.V. and Roalson, E.H. 2019. A well-known “mussambê” is a new species of Tarenaya (Cleomaceae) from South America. Syst. Bot. 44: 686–691. doi: 036364419X15620113920743

Soares Neto, R. L., Thomas, W. W., Vasconcellos Barbosa, M. R., and Roalson, E. H. (2020). Diversification of New World Cleomaceae with emphasis on *Tarenaya* and the description of *Iltisiella* , a new genus. *Taxon* 69, 321–336. doi:10.1002/tax.12231.

Southgate, D. (2009). Population Growth, Increases in Agricultural Production and Trends in Food Prices. *Electron. J. Sustain. Dev.*, 1. Available at: www.ejsd.org [Accessed August 2, 2021].

Stevens, P. F. Angiosperm Phylogeny Website. Version 14, July 2017 [and more or less continuously updated since]. Available at: http://www.mobot.org/mobot/research/apweb/orders/brassicalesweb.htm [Accessed August 2, 2021].

Thulin, M., and Roalson, E. H. (2017). Resurrection of the Genus *Rorida* (Cleomaceae), a Distinctive Old World Segregate of *Cleome*. *Syst. Bot.* 42, 569–577. doi:10.1600/036364417X695989.

U, S., and AP, Weber. (2016). The Road to C4 Photosynthesis: Evolution of a Complex Trait via Intermediary States. *Plant Cell Physiol.* 57. doi:10.1093/PCP/PCW009.

Valin, H., Sands, R. D., van der Mensbrugghe, D., Nelson, G. C., Ahammad, H., Blanc, E., et al. (2014). The future of food demand: Understanding differences in global economic models. *Agric. Econ. (United Kingdom)* 45, 51–67. doi:10.1111/AGEC.12089/ABSTRACT.

Voznesenskaya, E. V., Koteyeva, N. K., Chuong, S. D. X., Ivanova, A. N., Barroca, J., Craven, L. A., et al. (2007). Physiological, anatomical and biochemical characterisation of photosynthetic types in genus Cleome (Cleomaceae). *Funct. Plant Biol.* 34, 247. doi:10.1071/FP06287.

World Population Prospects: The 2017 Revision | Multimedia Library - United Nations Department of Economic and Social Affairs Available at: https://www.un.org/development/desa/publications/world-population-prospects-the-2017-revision.html [Accessed August 2, 2021].

CHAPTER 1

Brazil as a Center of *Tarenaya* (Cleomaceae) Diversity

Daniele F. Parma[1], Marcelo G. M. V. Vaz[1], Wagner L. Araújo[1], M. Eric Schranz[2], Andreas P. M. Weber[3] & Adriano Nunes-Nesi[1]*

[1] Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.
[2] Biosystematics Group, Wageningen University, 6708 PB Wageningen, The Netherlands.
[3] Institute for Plant Biochemistry Biology, Cluster of Excellence on Plant Science CEPLAS, Heinrich Heine University, Düsseldorf, Germany.

*Corresponding author: nunesnesi@gmail.com

ABSTRACT

In recent years, members of the family Cleomaceae have been analysed by means of morphological and molecular (phylogenetic) data, aiming to circumscript well-defined morphological groups at an evolutionary level. In this sense, currently, this family harbores 26 genera and about 270 species. Here, the phylogenetic relationships among members of *Tarenaya* and other Cleomaceae were evaluated using five molecular markers, including nuclear (ITS), chloroplastidic (*mat*K, *ndh*F and ycF) and mithocondrial (rps), as well as morphological and ecological data. These analyses revealed that the series of *Tarenaya,* Spinosa, Hemiscola, Parviflora and Rosea are monophyletic. In addition, it was shown that 26% of the Brazilian species of *Tarenaya* are endangered, 4% are in the vulnerable category and 70% may be in the least concern category. These latter results demonstrate the importance of studies with the group to propose strategies aiming their preservation.

Keywords: Brassicales, Cleomaceae, molecular phylogeny, morphology, tropical biomes.

INTRODUCTION

The morphological study of organisms is based on the description and analysis of their shape. Thus, morphology is one of the oldest areas and was widely used by naturalists to help us to name/identify the species and understand how phenotypes evolved (Wanninger, 2015). Although very important in the field of sciences, morphology has been losing ground to more modern techniques. In this sense, technological advances in molecular biology, together with the drastic reduction in sequencing costs, quickly resulted in an explosion of data useful for phylogenetic and developmental analyzes, mainly from the 1980s. In this regard, as more genes and genomes are sequenced, the possibility that thousands of informative molecular characters in evolution can be used in a given phylogenetic problem is quickly becoming a reality (Wiens, 2004). Based on this premise, there are some which advocate that only the use of molecular data is enough to describe and reclassify taxa (*e.g.* Sytsma et al., 1991; Turbeville et al., 1992). Anyway, using such methods for species delimitation and final taxonomic implication is not without problems and must be utilised with caution (DeSalle et al., 2005), especially when use a single-gene approach (Chase et al., 2005). In this case, integrative taxonomy, combining molecular, morphological, ecological and geographical data, is seen as the best route to identify, describe and classify taxa reliably (Puillandre et al., 2012).

However, it should be noted that with some exceptions, new species or phyletic groups are usually discovered, delimited and described using morphological data, mainly diagnostic characters (*e.g.* Soares-Neto et al. 2019; Parma et al., 2016; Soares-Neto et al., 2017; Pianissola et al., 2018). Nevertheless, the best methodology for delimiting species or groups using morphological data remains totally unresolved, and these different methods can give very different species limits for the same morphological data (*e.g.* Spinosa complex in the Cleomaceae family - see Soares-Neto et al. 2019). In this context, although this allows formal descriptions of taxonomic units, generally morphological data alone is not sufficient for a solid understanding of how organisms can be related to each other. For that, in our current scenario it would be indicated to use molecular phylogeny as a backbone to reconstruct the main evolutionary events and basic patterns of certain groups together with the morphology. In this sense, we can mention the work of molecular phylogeny with Cleomaceae (Feodorova et al., 2010; Patchell et al., 2014) that served as a framework for several taxonomic changes in the family that came soon after (*e.g.* Roalson et al.2015; Thulin and Roalson, 2017; Soares-neto et al., 2018).

Accordingly, with the advent of molecular phylogeny, several taxonomic changes have been proposed for different groups of organisms, which was not different for Cleomaceae. The family with 26 genera and about 270 species (Bayat et al., 2018; Soares Neto et al., 2020) has been the subject of several taxonomy and phylogeny studies (*e.g.* Iltis, 1957, 1959; Hall et al., 2002; Sanchez-Acebo, 2005; Inda et al., 2008; Feodorova et al., 2010; Iltis and Cochrane, 2014). Patchell et al. (2014), studying Cleomaceae, developed a molecular phylogeny using five markers (ITS, *mat*k, *yc*f, *ndh*F, rps3) from three genomes (nuclear, chloroplast and mitochondrial) with the main objective of understanding the *Cleome* s.s. Based on phylogenetic data, they were able to separate Cleomaceae into 15 clusters. Since then, many studies of taxonomy and systematics of Cleomaceae have used the work of Patchell et al. (2014) to validate proposed taxonomic changes with only morphological characteristics (*e.g.* Roalson et al. 2015; Thulin and Roalson, 2017; Sores-Neto et al., 2018). Accordingly, the main morphological characteristics used to separate the species in Cleomaceae are: growth form, pubescence, leaflet shape, petal size and stamen number.

Brazil can be considered a diversity center of Cleomaceae, with a record of 34 species; *Tarenaya*, with 23 species, is the most diverse genus of the family. It should be noted that *Tarenaya* comprises four mainly groupings: (i) Spinosa, (ii) Hemiscola, (iii) Parviflora and (iv) Rosea, and until recently the genus was considered polyphyletic (Inda et al., 2008; Feodorova et al., 2010; Patchell et al., 2014). However, some authors (*e.g.* Feodorova et al., 2010; Patchell et al., 2014) have proposed the hypothesis that polyphilia of *Tarenaya* might be resolved by including *Hemiscola* within *Tarenaya* or by breaking it into more genera (possible according to the series of *Tarenaya*). However, morphological characters associated with these clades need to be explored in more detail in order to determine the best appoach (Feodorova et al., 2010). Thus, Soares-Neto et al., (2018) made 25 new combinations for the genus *Tarenaya*, so many species that were previously recognized by *Cleome* or *Hemiscola* were transferred to the genus *Tarenaya*. It is noteworthy that this work was carried out based on morphology, considering mainly the synapomorphies of *Tarenaya* (spines and prickles of epidermal origin, seeds with an aril and a large cleft). After this taxonomic modifications, *Tarenaya* became a monophyletic group with around 37 species (Flora of Brazil 2021; Soares Neto et al., 2020).

Although most of the series of *Tarenaya* share synapomorphies of the genus (except *T. siliculifera*), morphological differences such as floral morphology, pollination strategies, seed characters and putative chromosome numbers (Iltis et al., 2007; Inda et al., 2008) stand out in comparison to the similarities. In this context, this study addresses the diversity of *Tarenaya* species occurring in Brazil, taking into account the molecular phylogeny, morphology and ecological data. For that, molecular phylogeny based on five markers ITS (nuclear), *mat*K, *ndh*F and ycF (chloroplast) and rps3 (mithocondria). In addition, we present the classification of the species in terms of conservation status according to the criteria adopted by the IUCN (2019).

MATERIAL AND METHODS

Taxon sampling

Sequences were obtained through botanical material of Cleomaceae from different Brazilian states, in addition to the sequences deposited at the NCBI, previously used in other works with the family (Feodorova et al., 2010; Patchell et al., 2014; Barrett et al., 2017; Soares Neto et al., 2020). More than one accession of species with broad distributions and/or with morphological variability were included whenever possible. In this way, 55 samples representing 39 species belonging to *Tarenaya* and genera phylogenetically close to it were sampled, *Cleoserrata, Dactylaena, Melidiscus, Physostemon*. Sequences of *Arabidopsis* were included as outgroup to molecular phylogeny.

Additionally, some Brazilian herbaria were also visited by the authors, such as BHCB, CESJ, HUEFS, CEPLAC, FLOR, MBM, UPCB and VIC (acronyms according to Thiers, 2018) in an effort to locate additional specimens for analysis and to a better understand of their distribution. In addition, consultation of available exsiccates on virtual herbarium was carried out (*e.g.* specieslink plataform). Information in *opus principes* (original protologues) was used to determine the specimens and for all names treated here.

DNA extraction, amplification and sequencing

Total DNA was extracted from fresh samples as well as from dried-herbarium material, using a modified CTAB-based method (Doyle et al., 1990) (Table S1). Five genetic regions (molecular markers) were investigated: three from the chloroplast – cpDNA (*ndhF*, *matk, ycF*), one from nucleus – nDNA (ITS), and one from mitochondria – mDNA (rps3) considered as phylogenetically informative at the inter- and infrageneric levels in the Cleomaceae family (Patchell et al., 2014; Soares-Neto et al. 2020). The segments of interest were amplified by means of polymerase chain reaction (PCR), which were implemented with a total volume of 50 μL: 0.5 ρmol of each oligonucleotide primer, 0.2 mM of each dNTP, 2.0 mM $MgCl_2$, 1 × PCR buffer (20 mM Tris-HCl (pH 8.4), 50 mM KCl), and 1.5 U Platinum Taq DNA polymerase (Invitrogen Life Technologies, Carlsbad, CA, USA). Primers used in this study were based on previous studies (Drew et al. 2011; White, 1990; Olmstead et al., 1993; Hall et al., 2002; Beilstein et al., 2006; Davis et al., 2006; Wurdack and Davis, 2009). The thermal cycling conditions were carried out according to Patchell et al., (2014). PCR products were visualized using 1 % (agar – *w*/*v* on TAE buffer) gel electrophoreses and then cleaned with Wizard SV Gel and PCR Clean-Up System PCR purification columns (Promega). The generated fragments were purified and submitted to DNA sequencing at Mylleus (https://store.myleus.com/). The sequenced fragments were assembled into one contig using the software Sequencer 4.1 (GeneCodes, Ann Arbor, Michigan). The aligned length of the data matrix including *mat*K (1,423 bp), *ndh*F (1,477 bp), ycf (1,044 bp), rps3 (1,272 bp), and ITS (1,191 bp) was combined (concatenated) for a total length of 6,411 bp for the 55 taxa included in this dataset.

Phylogenetic analyses

Phylogenetic inferences were carried out, based on the Maximum Likelihood (ML) and Bayesian inference (BI) methods (chloroplast, mitochondrial, and nuclear ribosomal). The ML analysis were carried out using RAxML version 8.2.10 and node support was obtained using the rapid bootstrap algorithm with 1,000 replicates (Stamatakis, 2014). The search for the best-scoring ML tree was performed with the general time-reversible model of DNA sequence evolution with gamma-distributed rate heterogeneity (GTRGAMMA model). The BI analyzes was performed according to the best fit evolutionary model which was selected by the Akaike information criterion (AIC) (Posada and Buckley, 2004) using MrModeltest 2.2 (Nylander, 2004). The Markov chain Monte-Carlo algorithm was executed with four runs with 50 million generations each and sampling every 5000 generations. The first 25% of the trees were discarded as burn-in, the remaining trees were used to construct the majority rule consensus tree, and then the PP for each node was calculated.

RESULTS AND DISCUSSION

Phylogenetic analysis

Phylogenetic reconstruction based on genetic sequences of the nuclear marker - ITS, chloroplast - *mat*K, *ndh*F and ycF - and mitochondrial - rps3, demonstrated that the genera *Melidiscus, Cleoserrata, Dactylaena, Tarenaya* and *Physostemon* were monophyletic, with the first three positioned as a sister group to *Tarenaya*. In *Tarenaya*, the series are monophyletic, Spinosa (BS: 84, PP: 1), Hemiscola (BS: 86, PP: 1), Parviflora (BS: 70, PP: 1) and Rosea (BS: - , PP: 0.96), as well as *Tarenaya* (BS: 98, PP: 1). Previous studies of molecular phylogeny that included members of *Tarenaya* showed that the series were polyphyletic (Sanchez-Acebo, 2005; Inda et al., 2008; Feodorova et al., 2010; Patchell et al., 2014; Parma et al. *submitted*). In the phylogeny ML with the combined all markers, the Rosea series proved to be polyphyletic, since the group Siliculosae, formed by the species *T. siliculifera*, was not grouped with the other species in the series Rosea (Supplementary Fig. 1). This is an interesting fact, as *T. siliculifera* is a morphologically different species from the other *Tarenaya* species, including the absence of spines at the base of the petiole, and different type of fruit, capsule with only four seeds (all other species of the genus have fruit elongated with at least 25 seeds). However, with a larger sample of taxa, Soares Neto et al., (2020) and the present study demonstrated that are monophyletic groups in BI. However, it is important to emphasize that here, the phylogenetic inferences are also supported by morphological data, as described below. It should be noted that not all species of *Tarenaya* were analyzed by means of molecular phylogeny. These specimes are difficult to collect, having only old records, which also hampers the obtaining of sequences using herbaria materials. More important, we do not know how they were collected and stored before being dropped in herbaria.

The tree combined with all markers showed the same morphology as the phylogenetic tree based only on the nuclear marker ITS in Inference Baysian (Fig. 1). Thus, the Parviflora series (PP: 1) is the outermost group of the *Tarenaya* genus, followed by Hemiscola (PP: 0.99) and later Rosea (PP: 0.95) and Spinosa (PP: 1). Although, both trees have the same topology, the phylogeny combined to five marks presented the highest support value, mainly among species. In turn, in the tree combined with the chloroplast markers, the series are presented as a comb, in addition to *T. siliculifera* (Siliculosae) grouped with the species of the Parviflora series. Thus, the marker that best portrays the evolutionary history of the group is the ITS or a combination of all.

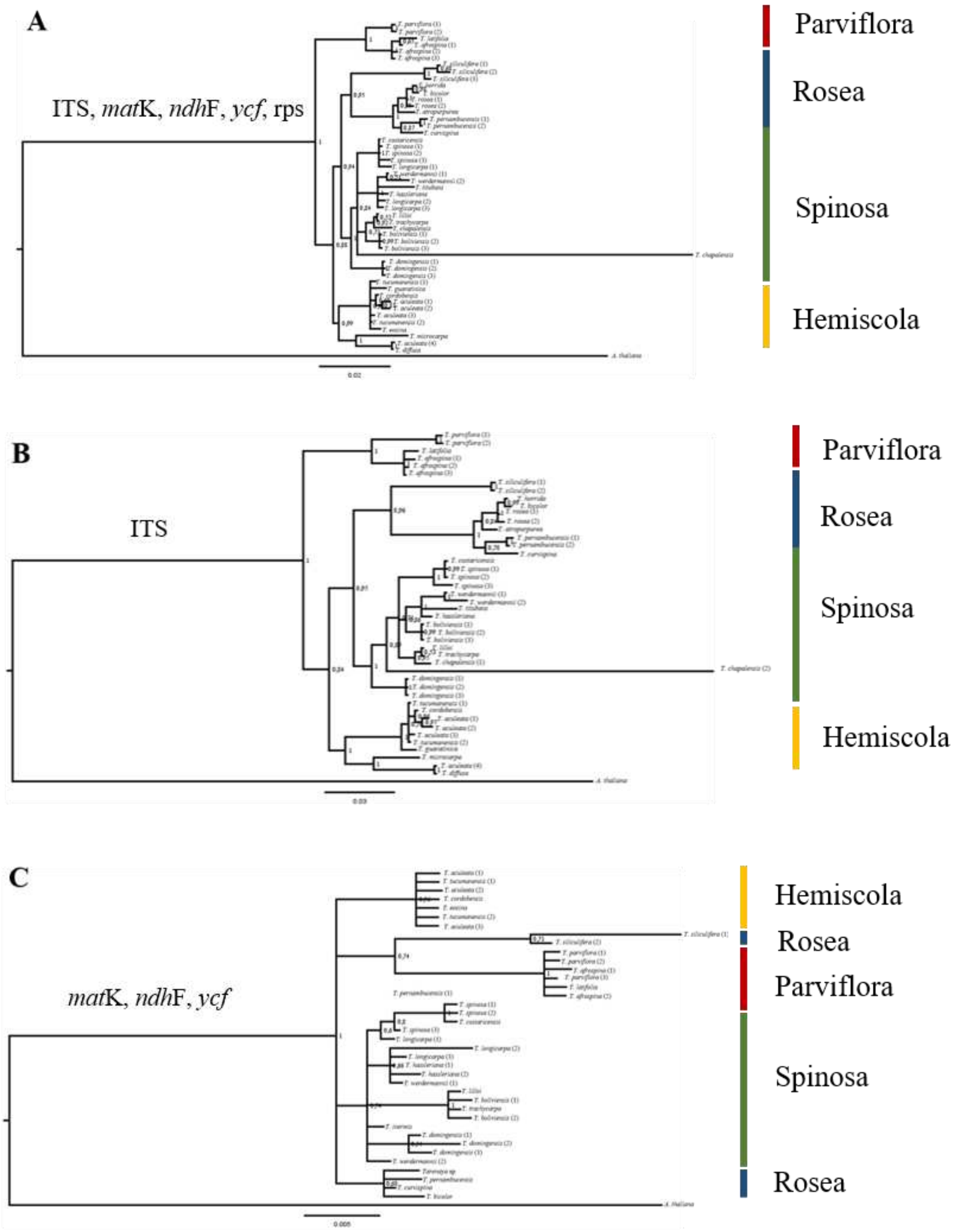


Figure 1. Phylogenetic relationships of Cleomaceae species inferred from Inference Baysian. Number at nodes reflect posterior probability. A. Phylogenetic relationships based on combined sequences of ITS, *mat*K, *ndh*F, ycF and rps, bar: 0.02 subtitutions per nucleotide. B. Phylogenetic relationships based on ITS, bar: 0.02 subtitutions per nucleotide. C. Phylogenetic relationships based on combined sequences of *mat*K, *ndh*F, ycF, bar: 0.005 subtitutions per nucleotide.

Morphological characterization of groups

Tarenaya is a New World genus segregated from *Cleome* and was reestablished by Iltis and Cochrane (2007) based on seed characters and chromosome number. However, *Tarenaya* was recognized as not monophyletic because species of *Hemiscola* were nested within this genus, and the inclusion of these species into a broad concept of *Tarenaya* was proposed (Patchell et al., 2014; Arana and Oggero, 2016; Soares-Neto et al., 2018). Species of these genera share the pair of spines at the base of the petioles and the seeds with a membrane in the opening gap. Thus, Soares-Neto et al., (2018) proposed to expand the genus *Tarenaya* to include the species of *Hemiscola* and the other species of *Cleome* sect. *Tarenaya* not yet transferred to *Tarenaya,* forming a monophyletic group.

Tarenaya species sampled for molecular phylogeny are recognized as belonging to four series mainly (Hemiscola, Parviflora, Spinosa and Rosea), which are also considered monophyletic with high support value. With this new definition, *Tarenaya* comprises 37 species (Soares-Neto et al., 2018; Soares Neto et al., 2020). In Brazil, genera diversity center, there are 23 species (Flora of Brazil 2021; Soares-Neto et al., 2019; Soares Neto et al., 2020), from which 18 are endemic. Although Cleomaceae (35Mya) and, mainly, *Tarenaya* (17 Mya) (Soares-Neto et al., 2020) have a recent speciation, species in this group have a great morphological diversity, genetic and physiology (see Parma et al. *submitted*). Furthemore, *Tarenaya* still has the most recent common ancestor in the Chacoan and Atlantic Forest geographic areas (Soares-Neto et al., 2020). Thus, Rosea would be the group that emerged first (17Mya), followed by the other groups (15Mya) (Soares-Neto et al., 2020). Perhaps, this justifies the fact that Spinosa is the most specious group of the genus, however, in the country (Atlantic Forest) the Rosea group predominates in number of species (Table 1).

In addition to monophylia, the series form groups that can be morphologically recognized (see Table 1). In fact, two out of four lineages here represent morphologically distinct clusters (Spinosa and Hemiscola) and are thus relatively straight-forward to be separated and recognized as independent taxa. The other two clusters (Parviflora and Rosea) are susceptible to separation, but some of their morphological traits overlap. However, despite this morphological overlap, these two monophyletic groups are not closely related considering the phylogenetic reconstruction (Fig. 1). In addition, it is important to note that the species of the Parviflora and Rosea groups occur in different areas. In this sense, Rosea is formed by a group of Brazilian endemic species that are distributed mainly in the south and southeast regions and some states in the northeast (mainly Bahia state). While members of Parviflora have a wider distribution in other countries and with few representatives in Brazil, often restricted to the northern region (except *T. parviflora*).

Table 1. Brazilian species of *Tarenaya* distributed in the corresponding series

Spinosa	**Hemiscola**	**Parviflora**	**Rosea**
T. hassleriana	*T. aculeata*	*T. latifolia*	*T. bernadeteana*
T. longicarpa	*T. atropurpurea*	*T. parviflora*	*T. bicolor*
T. titubans	*T. crenopetala*	*T. psoraleifolia*	*T. curvispina*
T. trachycarpa	*T. diffusa*	*T. virens*	*T. horrida*
	T. eosina		*T. inermis*
	T. microcarpa		*T. pernambucensis*
			T. regnelli
			T. rosea
			T. siliculifera

***Tarenaya* series**

Spinosa – The Spinosa group has sub-shrub to shrubby habit, leaves usually with 5-7 leaflets, large and showy flowers arranged in a terminal elongated inflorescence, elongated fruits (Fig. 2, Table 2). In addition, morphological features as pollen (spinolose pollen grains) and seeds morphology (cleft covered by a thin membrane) support Spinosa as an independent group (Sanchez-Acebo, 2005; Inda et al., 2008). Spinosa encompasses about 10 species, widely distributed in the American continent (central and south), however in Brazil has only four species (Table 1), with wide distribution in the territory.

Parviflora – The Parviflora group has sub-shrub habit, leaves usually with 3-5 leaflets, small, delicate flowers arranged in an elongated terminal inflorescence, elongated fruits (Fig. 2, Table 2).The Parviflora group differ slightly from Spinosa by having an obovoid to nearly spherical cleft cavity that is relatively large in longitudinal section. However, all members of the Spinosa and Parviflora clusters have smooth or scurfy seed surface. Parviflora species have a distribution in the American continent, with one species in

Table 2. Morphological features used to recognize each of the four series of *Tarenaya*

	Spinosa	**Hemiscola**	**Parviflora**	**Rosea**
Habit	shrub	herb	sub-shrub	herb
Petiole	3.5–8 cm	1.3–8.3 cm	3.5–13 cm	5–13 cm
Leaflet number	5-7	3-5	3-5	3-7
Leaflet shape	elliptic to oblanceolate-elliptic, basally attenuate, apically acute to acuminate, margin serrulateciliate	elliptic to oblanceolate, basally cuneate to attenuate, apically acute to acuminate or obtuse, margin entire to slightly serrulate-ciliate	elliptic to widely elliptic, elliptic-lanceolate, basally attenuate to cuneate, apically acuminate to long acuminate, margin sinuose-ciliate to serrulate-ciliat	elliptic to widely elliptic, elliptic-lanceolate, basally attenuate, apically acuminate to long acuminate, margin entire or serrulate to serrulate-ciliate
Pedicel	15–30 mm	8–15 mm	10–20 mm	18–25 mm
Petal: Claw length x width	7–15 × 3–5 mm	2–6 × 2.5–4 mm	3–4 × 1.5 mm	8–10 × 2–4 mm
Stamen length	21–33 mm	4.5–11 mm	10–16 mm	19–25 mm
Fruit length	50–120 mm	13–30 mm	47–80 mm	50–87 mm
Seed length	1.5–2 × 1–1.2 mm	2–2.5 × 1–1.5 mm	1.5–2.2 × 1 mm	1.5–2 × 1 mm
Seed morphology	longitudinally striate, with few transversely ridged; cleft covered by a membrane.	longitudinally striate, with transverse ridges covered by simple hair; cleft covered by a membrane and an aril attaching the claws.	longitudinally striate and finely transversely ridged; cleft covered by a membrane.	longitudinally striate, conspicuously transversely ridged; cleft covered by a membrane.

Africa (*T. afrospina*). With the exception of *T. parviflora*, which has a wide occurence, followed by *T. psoraleofolia*, the other species are restricted to some region and/or country (*e.g. T. torticarpa* in Venezuela). In Brazil, there are few representatives of this group (four species) that are often restricted to the North, such as *T. latifolia, T. parviflora, T. psoraleofolia* and *T. virens.*

Rosea – The Rosea group has herb habit, leaves usually with 3-7 leaflets, medium sized flowers, compared to other groups, arranged in an elongated terminal inflorescence, also medium sized fruit. This group comprises only species with distribution in Brazil (Flora of Brazil 2021). Some of these have distribution restricted to some region and/or state (*e.g. T. inermis, T. regnelli, T. bernadeteana*). It is worth mentioning that some of the current species were considered as a subspecies/variety of *T. rosea*, such as, *T. bicolor, T. pernambucensis and T. inermis*.

Hemiscola - The Hemiscola group has herbaceous habit, smaller size, small and delicate flowers often arranged in axillary, sometimes terminal inflorescences, smaller fruits and seeds with aryl, which sets it apart from Spinosa, for example. Although the two groups share the same chromosomal number (x = 8-10 - Inda et al., 2008), it does not have the same genome size, in addition to presenting different physiological and biochemical responses (see Parma et al. *submitted*).

The Hemiscola group was genus *Hemiscola* which was segregated from *Cleome s.l.* by Rafinesque (1838) and includes only *T. aculeata*. Later, *Cleome diffusa* was inserted in the genus *Hemiscola* (*H. difusa* (Banks ex DC.) Iltis) (Iltis et al., 2007). However, adopting a broader view of the genus, characterized by white to cream flowers in short, bracteate inflorescences or solitary in the axils of leaves and by horseshoe-shaped seeds that have the mouth of the cleft covered by a thin membrane, the tip of the radicular claw expanded into a large, white or ivory, funicular aril, and the forehead transverse-wrinkled or irregularly tubercled, the ridges or excrescences sparsely pubescent on their crests with minute clusters of caducous white hairs, the series could encompass less than 10 spiny species (and the rare unarmed forms – *e.g. T. crenopetala*) (Iltis et al., 2007). The species included in this clade are distributed throughout South America, except *T. aculeata* which has a world-wide distribution. In addition, it is worth noting that some species that currently belong to the *Hemiscola* cluster were recognized as subspecies/morphospecies of *T. aculeata* (*e.g. T. diffusa*, *T. guaranitica, T. cordobensis*).

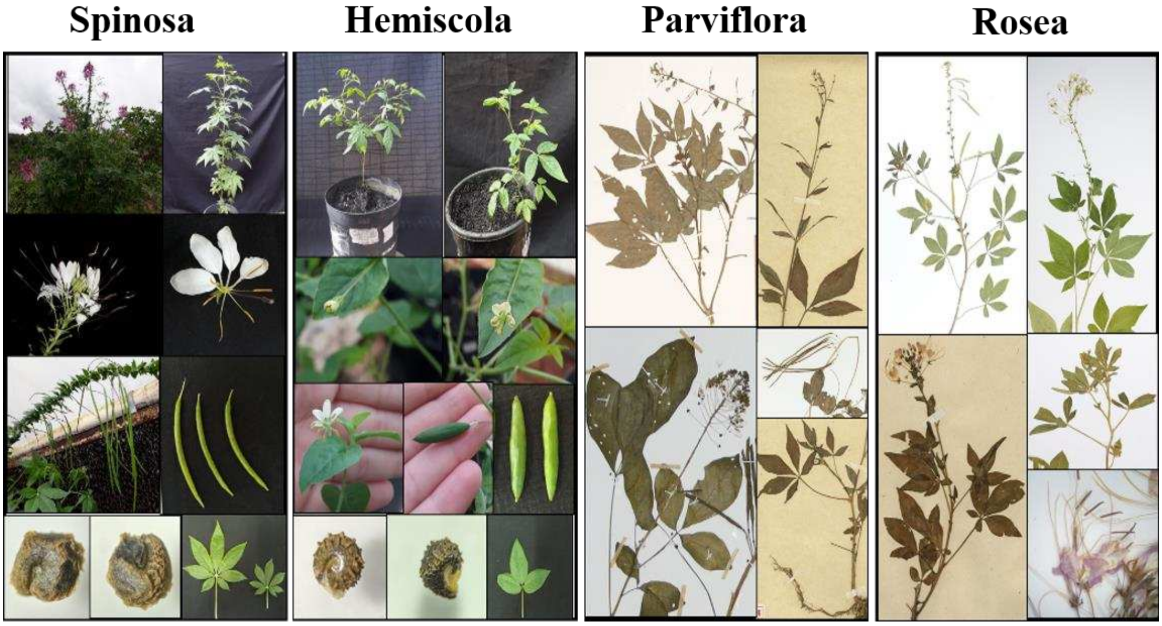


Figure 2. Pictures with the main characteristics of representatives of *Tarenaya* (Spinosa, Hemiscola, Parviflora and Rosea).

Conservation status of *Tarenaya* species

The species *T. bernadeteana, T. curvispina, T. inermis, T. regnelli, T. siliculifera* and *T. crenopetala* are classified as endangered (EN) by IUCN (B2b) criteria due to the low number of locations (less than or equal to 5) in which species can be found, with old records. In turn, *T. titubans* can be classified as vulnerable (VU) by the criteria of IUCN (B2b) due to the low number of locations (less than or equal to 10) in which the species can be found, in addition to the records being old. As for the others 16 species of the group *Tarenaya*, they can be classified as less concern (LC) because they have a wide distribution and/or several old and recent records.

CONCLUSIONS

In general, despite the *Tarenaya* series being monophyletic, the genus is really a very complex group and some species are in the boundary, having traits of different series, hampering an overall work to solve their systematic. In addition, it was shown that 22% of the *Tarenaya s.l.* species are endangered, 13% are in the vulnerable category and 65% may be in the least concern category. These latter results demonstrate the importance of studies with the group to propose strategies aimed at their preservation.

ACKNOWLEDGMENT

This work was supported by funding from the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) (grant number 424024/2018-7), the Foundation for Research Assistance of the Minas Gerais State (Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG) (grant number APQ-00528-18). Research fellowships granted by CNPq to ANN and WLA as well as the scholarship granted by the Coordination of Improvement of Higher Education Personnel (CAPES) to DFP are also gratefully acknowledged. MGMVV was supported by scholarships from CAPES (PNPD-1638006) and FAPEMIG (BPD-00514-14). APMW acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2048/1 – project ID 390686111.

REFERENCES

Arana, M. D., and Oggero, A. J. (2016). New combinations in Tarenaya (Cleomaceae) for the Argentinian flora. 267, 162–164.

Barrett, R. L., Roalson, E. H., Ottewell, K., Byrne, M., Govindwar, S. P., Yadav, S. R., et al. (2017). Resolving Generic Boundaries in Indian-Australasian Cleomaceae: Circumscription of *Areocleome* , *Arivela* , and *Corynandra* as Distinct Genera. *Syst. Bot.* 42, 694–708. doi:10.1600/036364417X696401.

Bayat, S., Schranz, M. E., Roalson, E. H., and Hall, J. C. (2018). Lessons from Cleomaceae , the Sister of Crucifers. *Trends Plant Sci.* xx, 1–14. doi:10.1016/j.tplants.2018.06.010.

Beilstein, M. A., Al-Shehbaz, I. A., and Kellogg, E. A. (2006). Brassicaceae phylogeny and trichome evolution. *Am. J. Bot.* 93, 607–619. doi:10.3732/ajb.93.4.607.

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N., et al. (2005). Land plants and DNA barcodes: short-term and long-term goals. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1889–1895. doi:10.1098/rstb.2005.1720.

Davis, C. C., Latvis, M., Nickrent, D. L., Wurdack, K. J., and Baum, D. A. (2006). Floral Gigantism in Rafflesiaceae. 31242.

DeSalle, R., Egan, M. G., and Siddall, M. (2005). The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1905–1916. doi:10.1098/rstb.2005.1722.

Doyle, J., Doyle, J. J., et al. (1990). Isolation of plant DNA from fresh tissue.

Drew, B. T., Sytsma, K. J., Drew, B. T., and Sytsma, K. J. Testing the Monophyly and Placement of Lepechinia in the Tribe Mentheae (Lamiaceae) Testing the Monophyly and Placement of Lepechinia in the Tribe Mentheae (Lamiaceae). 36, 1038–1049. doi:10.1600/036364411X605047.

Feodorova, T. A., Voznesenskaya, E. V, Edwards, G. E., Roalson, E. H., Feodorova, T. A., Voznesenskaya, E. V, et al. (2010). Biogeographic Patterns of Diversification and the Origins of C4 in Cleome (Cleomaceae). 35, 811–826. doi:10.1600/036364410X539880.

Flora of Brazil - Cleomaceae Bercht. & J.Presl Available at: http://floradobrasil.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/BemVindoConsulta

PublicaConsultar.do?invalidatePageControlCounter=1&idsFilhosAlgas=%5B2%5D&idsFilhosFungos=%5B1%2C10%2C11%5D&lingua=&grupo=5&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=Cleomaceae+Bercht.+%26+J.Presl&formaVida=null&substrato=null&ocorreBrasil=QUALQUER&ocorrencia=OCORRE&endemismo=TODOS&origem=TODOS&regiao=QUALQUER&estado=QUALQUER&ilhaOceanica=32767&domFitogeograficos=QUALQUER&bacia=QUALQUER&vegetacao=TODOS&mostrarAte=SUBESP_VAR&opcoesBusca=TODOS_OS_NOMES&loginUsuario=Visitante&senhaUsuario=&contexto=consulta-publica [Accessed February 17, 2021].

Guidelines for Using the IUCN Red List Categories and Criteria ™ (2017). Available at: The Iucn Red List Of Threatened Specieshttp://www.iucnredlist.org/documents/RedListGuidelines.pdf. [Accessed may 24, 2021].

Hall, J. C. (2008). Systematics of Capparaceae and Cleomaceae : an evaluation of the generic delimitations of Capparis and Cleome using plastid DNA sequence data 1. 696, 682–696. doi:10.1139/B08-026.

Hall, J. C., Sytsma, K. J., and Iltis, H. H. (2002a). Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data; Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. doi:10.3732/ajb.89.11.1826.

Hall, J. C., Sytsma, K. J., and Iltis, H. H. (2002b). Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. *Am. J. Bot.* 89, 1826–1842. doi:10.3732/ajb.89.11.1826.

Iltis, H. H. (1957). Studies in the Capparidaceae. III. Evolution and Phylogeny of the Western North American Cleomoideae. *Ann. Missouri Bot. Gard.* 44, 77. doi:10.2307/2394679.

Iltis, H. H. (1959). mens above the basic number of 6 , therefore , seems to be a derived rather than and Ranmanissa have many stamens , in some species of the latter often close to.

Iltis, H. H., and Cochrane, T. S. (2014). Studies in the Cleomaceae VI : A New Genus and Sixteen New Combinations for the Flora Mesoamericana Studies in the Cleomaceae VI : A New Genus and Sixteen New Combinations for the Flora Mesoamericana. doi:10.3417/2013017.

Iltis, H. H., Cochrane, T. S., Iltis, H. H., and Cochrane, T. S. (2007). Studies in the Cleomaceae V : A New Genus and Ten New Combinations for the Flora of North America Studies in the Cleomaceae V : A New Genus and Ten New Combinations for the Flora of North America. 17, 447–451.

Inda, L. A., Torrecilla, Æ. P., Catala, Æ. P., and Ruiz-zapata, T. (2008). Phylogeny of Cleome L . and its close relatives Podandrogyne Ducke and Polanisia Raf . (Cleomoideae , Cleomaceae) based on analysis of nuclear ITS sequences and morphology. 111–126. doi:10.1007/s00606-008-0026-y.

Luciano, R., Neto, S., Thomas, W. W., Regina, M., Barbosa, D. V., Roalson, E. H., et al. (2018). New combinations and taxonomic notes for Tarenaya (Cleomaceae). 1–6. doi:10.1590/0102-33062017abb0417.

Molecula, S. (1990). White , T . J ., T . D . Bruns , S . B . Lee , and J . W . Taylor . Amplification and direct sequencing of fungal ribosomal RNA Genes for ...

Neto, R. L. S., Regina, M., Barbosa, D. E. V., and Eric, H. (2017). Cleoserrata (Cleomaceae): taxonomic considerations and a new species. 324, 179–186.

Nylander (2004). MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Olmstead, R. G., Jennifer, A., Wolfe, K. H., and Biology, E. P. O. (1993). Update section Ninety extra nucleotide in ndhF gene of tobacco chloroplast DNA : a summary of revisions to the

1986 genome sequence. 1191–1193.

Parma, D. F., Pianissola, E. M., Vinícius-Silva, R., Clark, L. G., and Santos-Gonçalves, A. P. (2016). Two new species of Merostachys (Poaceae: Bambusoideae) from the Brazilian Atlantic forest. *Phytotaxa* 267, 211–218. doi:10.11646/phytotaxa.267.3.5.

Patchell, M. J., Roalson, E. H., and Hall, J. C. (2014). Resolved phylogeny of Cleomaceae based on all three genomes. 63, 315–328.

Pianissola, E. M., Parma, D. F., Santos-Gonçalves, A. P., and Clark, L. G. (2018). Two new species of Chusquea subg. Swallenochloa (Poaceae: Bambusoideae: Bambuseae) from Minas Gerais, Brazil, and complete description of C. caparaoensis. *Phytotaxa* 358, 235–250. doi:10.11646/phytotaxa.358.3.2.

Posada, D., and Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53, 793–808. doi:10.1080/10635150490522304.

Puillandre, N., Modica, M. V., Zhang, Y., Sirovich, L., Boisselier, M. C., Cruaud, C., et al. (2012). Large-scale species delimitation method for hyperdiverse groups. *Mol. Ecol.* 21, 2671–2691. doi:10.1111/j.1365-294X.2012.05559.x.

Raimundo, A., Soares, L., Thomas, W. W., De, M. R., Barbosa, V., Roalson, E. H., et al. A Well-Known " Mussambê " is a New Species of Tarenaya (Cleomaceae) from South America 44, 686–691. doi:10.1600/036364419X15620113920743.

Roalson, E. H., Hall, J. C. 2017.. New Generic Concepts for African Cleomaceae New Generic Concepts for African Cleomaceae. 42, 925–942. doi:10.1600/036364417X696393.

Sanchez-Acebo, L. (2005). A Phylogenetic Study of the New World Cleome (Brassicaceae, Cleomoideae) on JSTOR. *Ann. Missouri Bot. Gard.* 92, 179–201.

Soares Neto, R. L., Thomas, W. W., Vasconcellos Barbosa, M. R., and Roalson, E. H. (2020). Diversification of New World Cleomaceae with emphasis on *Tarenaya* and the description of *Iltisiella* , a new genus. *Taxon* 69, 321–336. doi:10.1002/tax.12231.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi:10.1093/bioinformatics/btu033.

Sytsma, K. J., Smith, J. F., and Berry, P. E. (1991). The Use of Chloroplast DNA to Assess Biogeography and Evolution of Morphology, Breeding Systems, and Flavonoids in Fuchsia sect, Skinnera (Onagraceae). *Syst. Bot.* 16, 257. doi:10.2307/2419278.

Thulin, M., and Roalson, E. H. (2017). Resurrection of the Genus *Rorida* (Cleomaceae), a Distinctive Old World Segregate of *Cleome*. *Syst. Bot.* 42, 569–577. doi:10.1600/036364417X695989.

Turbeville, J. M., Field, K. G., and Raff, R. A. (1992). Phylogenetic position of phylum Nemertini, inferred from 18S rRNA sequences: molecular data as a test of morphological character homology. *Mol. Biol. Evol.* 9, 235–249. doi:10.1093/oxfordjournals.molbev.a040716.

Wanninger, A. (2015). Morphology is dead â€" long live morphology! Integrating MorphoEvoDevo into molecular EvoDevo and phylogenomics. *Front. Ecol. Evol.* 3, 54. doi:10.3389/fevo.2015.00054.

Wiens, J. J. (2004). The Role of Morphological Data in Phylogeny Reconstruction. *Syst. Biol.* 53, 653–661. doi:10.1080/10635150490472959.

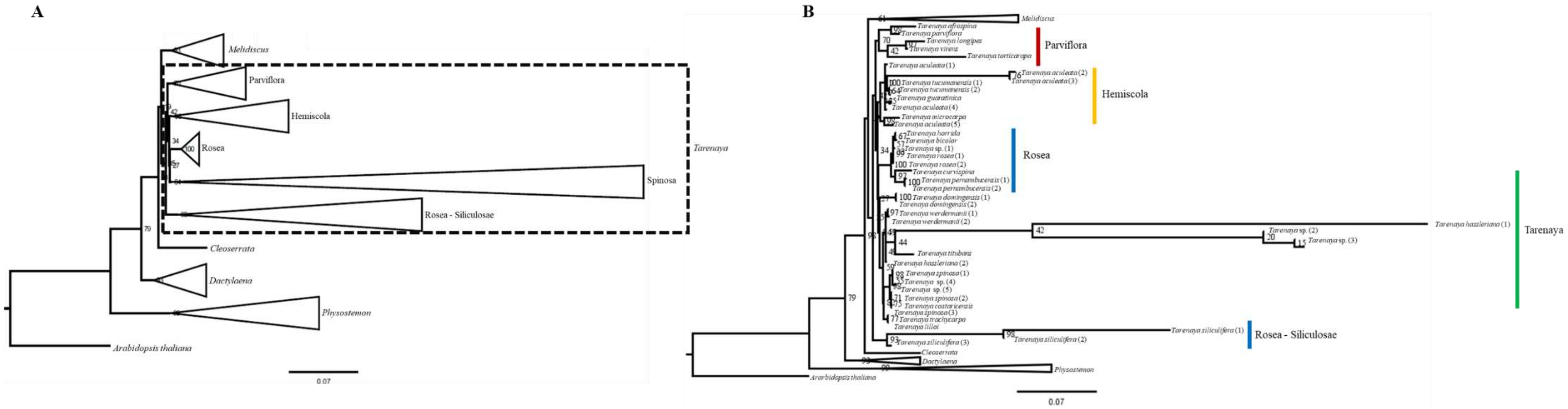
Wurdack, K. J., and Davis, C. C. (2009). Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm. 96, 1551–1570. doi:10.3732/ajb.0800207.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Species used in molecular phylogeny analyses and their respective number of identification – accession number at NCBI. *Species collected in this study.

Species	**ITS**	***mat*K**	***ndh*F**	**ycF**	**rps**
Arabidopsis thaliana	MG886682				
Cleoserrata speciosa	KT588831	KT588817			
Dactylaena micrantha	MN879425				
D. microphylla	HM044279	EU371809	EU373716		KF923229
D. pauciflora	MN879426	EU371810	EU373717		
Dactylaena sp. (1)	MN879423				
Dactylaena sp. (2)	MN879424				
Melidiscus giganteus (1)	HM044283	KF923135	KF923095	KF923037	
Melidiscus giganteus (2)*			MT094429	KF923066	
Physostemon lancolatum	KF923175				
P.rotundifolium	MN879431				
Tarenaya aculeata (1)	MN879435	EU371790	AY122382	KF923025	KF923190
T. aculeata (2)	MN879434				
T. aculeata (3)	MN879433				
T. aculeata (4)	MN879432				
T. aculeata (5)*	MT090703		MT094417		
T. afrospina	MN879437	KF923123			
T. bicolor	MN906009				
T. costaricensis	MN906010				
T. curvispina	MN879443				
T. domingensis (1)	MN879445	EU371793	AY122383	KF923033	KF923197
T. domingensis (2)	MN879444				
T. guaranitica	MN879446				
T. hassleriana (1)	DQ455791	AY491649		KF923039	
T. hassleriana (2)*	MT090705	MT094412	MT094420		
T. horrida	MN879447				
T. lilloi	MN879449				
*T. longicarpa**	MT090701	MT094410	MT094422		
T. longipes	MN879453				
T. microcarpa	MN879454				
T. parviflora	MN879455	EU371801		KF923053	KF923215
T. pernambucensis (1)	MN879458				
T. pernambucensis (2)	MN879457				
T. rosea (1)	MN879460		EU373710	KF923056	KF923218
T. rosea (2)	MN879459				
T. siliculifera (1)*	MT090717	MT094407	MT094418		
T. siliculifera (2)*			MT094428		
T. siliculifera (3)	HM044286	KF923155	KF923113	KF923059	KF923222
Tarenaya sp. (1)	MN879464				
Tarenaya sp. (2)	MN879452				
Tarenaya sp. (3)	MN879451				
Tarenaya sp. (4)	MN879450				
Tarenaya sp. (5)	DQ455798				
T. spinosa (1)	MN879463	JQ587203	DQ200029	KF923060	
T. spinosa (2)	MN879462	EU371805			
T. spinosa (3)	HM044296				
T. titubans	DQ455813				
T. torticarpa	DQ455810				

T. trachycarpa	HM044297	KF923158	KF923116		KF923224
T. tucumanensis (1)	MN879465	KF923159	KF923117	KF923061	KF923225
T. tucumanensis (2)	DQ455811				
T. virens	MN879466				
T. werdermannii (1)	MN879467	KF923163	KF923121		
T. werdermannii (2)	DQ455809				



Supplementary Figure 1. Phylogenetic relationships of Cleomaceae species inferred from Maximum Likelehood based on concatened sequences of ITS, *mat*K, *ndh*F, ycF and rps. Number at nodes reflect bootstrap values. Bar: 0.07 subtitutions per nucleotide. A. Relationship between the collapsed clusters of genus of Cleomaceae. B. Highlight for the phylogenetic relationship between species of *Tarenaya*.

CHAPTER 2

Proposal of *Neotarenaya siliculifera* gen. nov. *et* comb. nov.: A new generic unit erected from Brazilian specimens of *Tarenaya* (*Cleome*) *siliculifera*

Daniele F. Parma[1], Marcelo G. M. V. Vaz[1], Wagner L. Araújo[1], M. Eric Schranz[2], Andreas P. M. Weber[3] & Adriano Nunes-Nesi[1]*

[1]Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.
[2]Biosystematics Group, Wageningen University, 6708 PB Wageningen, The Netherlands.
[3]Institute for Plant Biochemistry Biology, Cluster of Excellence on Plant Science CEPLAS, Heinrich Heine University, Düsseldorf, Germany.

*Corresponding author:
Adriano Nunes-Nesi
Departamento de Biologia Vegetal,
Universidade Federal de Viçosa,
36570-900 Viçosa, Minas Gerais, Brazil
Phone: +55-31-3612-5357
Email: nunesnesi@ufv.br

ABSTRACT

Species of the family Cleomaceae have been recently analyzed by means of morphological and molecular (phylogenetic) data, aiming to circumscript well-defined groups at an evolutionary level. In this sense, currently 25 genera and about 270 species are recognized in the family. Among these taxa, the South American genus *Tarenaya* requires additional taxonomic revision. Here, we have examined a subset of 24 novel Brazilian *Tarenaya* morphospecies and evaluated their taxonomic/phylogenetic status, particularly *T. siliculifera*, which nomenclatural status has been already challenged previously. Here, the phylogenetic relationships among members of *Tarenaya* and other Cleomaceae were examined using three molecular markers, including one nuclear (ITS) and two chloroplastic (*matk* and *ndhF*), as well as morphological and ecological data. Our analyses revealed that the *Tarenaya* cluster/series, Spinosae, Hemiscola, Parviflorae and Cleorosea, are indeed polyphyletic. Molecular, morphological and ecological data strongly support further nomenclatural modifications for these taxa. Therefore, the species *Tarenaya siliculifera* is described here as a new monotypic genus, *Neotarenaya siliculifera* (voucher VIC 51801), under the provisions of the International Code of Nomenclature for algae, fungi, and plants. Our finding reinforces the importance of further collection and characterization of specimens sampled from tropical underexplored biomes for accurate taxonomic resolution and systematics.

Keywords: *Neotarenaya*, Cleomaceae, Molecular phylogeny, Morphology, Tropical biomes.

1. Introduction

Cleomaceae Berchtold & J. Presl is a botanical family with a worldwide distribution, belonging to the order Brassicales Bromhead and is the sister-family to the Brassicaceae (Iltis et al., 2011). Previously, Cleomaceae was classified as a subfamily of Capparaceae Juss (Hall et al., 2002). However, the monophyletic status of the families Brassicaceae, Cleomaceae, and Capparaceae, as well as their independent evolutionary origins, were demonstrated in molecular phylogenetic studies based on the chloroplastic markers (*matk*, *trn*K and *ndhF*) (Hall et al., 2002; Hall, 2008). In addition, the monophyly of these families was also confirmed by means of morphological characteristics (Iltis et al., 2011). Hence, the family Cleomaceae is recognized and distinguished as having the following morphological traits: palmately compound leaves, a racemose inflorescence, a gynophore present and monosymmetric flowers with the petals and stamens curved upwards (Stevens, 2001).

Considering previous studies, which were mainly focused on the molecular phylogeny of Cleomaceae, it is clear that the generic boundaries within the family are rather problematic (Feodorova et al., 2010; Hall, 2008; Hall et al., 2002; Inda et al., 2008; Patchell et al., 2014; Sanchez-Acebo, 2005). In recent years, several taxonomic changes have been proposed (Luciano et al., 2018; Roalson et al., 2015; Thulin and Roalson, 2017), mainly after the molecular phylogeny of Patchell et al., (2014), based on that recognized 15 clusters in *Cleome sensu lato* (hereafter, *Cleome* s.l.) (Feodorova et al., 2010; Patchell et al., 2014). However, it is also clear that Cleomaceae is in need of a modern systematic treatment. Several recent studies have proposed and recognized some genera by integrating morphological and molecular criteria for clades that are clearly monophyletic such as *Tarenaya*, *Cleomella* and *Rorida* (Luciano et al., 2018; Roalson et al., 2015; Thulin and Roalson, 2017). Currently, twenty-five genera and about 270 species are recognized (Bayat et al., 2018). However, this still includes several morphogenera that are not yet fully resolved.

Following the recognition of Cleomaceae as a family, significant efforts were devoted to understand the true circumscription of the genus *Cleome* s.l. which still remains polyphyletic (group Andean and *Cleome* s.s.). In this sense, different genera were used to understand/delimit their true circumscription (Barrett et al., 2017; Luciano et al., 2018; Roalson et al., 2015.; Thulin and Roalson, 2017). *Tarenaya* was the largest recognized section in *Cleome,* being characterized by seeds with two claws, mostly connected by a thin membrane and with a large cleft-cavity. These morphological traits have been recognized as putatively primitive and this section has a primary center of distribution in Southeast South America (Inda et al., 2008). Recently, this section was raised to the genus level as *Tarenaya* (Luciano et al., 2018). Thus, *Tarenaya* became one of the most diverse genera in number of species in the family Cleomaceae, harboring about 37 species (Luciano et al., 2018). In Brazil, 20 species have been recognized of which 17 are endemic (Flora of Brazil, 2020). Besides their clear molecular phylogenetic separation, the genus *Tarenaya* is easily distinguished from other genera by having spines and prickles of epidermal origin and seeds with an aril and a large cleft (Luciano et al., 2018a).

Molecular phylogenies based on ITS1 (nuclear marker) suggest that *Tarenaya*, in its present circumscription, is essentially monophyletic (Patchell et al., 2014). However, the current status of *T. siliculifera* (Eichler) Soares-Neto & Roalson (by Luciano et al., 2018) into the *Tarenaya* cluster was previously questioned (Patchell et al., 2014). *Tarenaya siliculifera* is an endemic species occurring in the “Cadeia do Espinhaço” of Brazil, a mountain range found in Atlantic uplands in the states of Minas Gerais and Bahia. According to Patchell et al (2014), *T. siliculifera* should be raised to the generic level based on molecular phylogenetic and morphological analysis. This species has small and delicate flowers, morphological characteristics that differentiate it from the other *Tarenaya* species. However, these traits are not restricted to *T. siliculifera* as they overlap with those for other species, such as *T. aculeata* (L.) Soares-Neto & Roalson, *T. diffusa* (Banks ex DC) Soares-Neto & Roalson and *T. microcarpa* (Ule) Soares-Neto & Roalson.

Taking into account (i) the ecological distribution of *Tarenaya siliculifera*, which is very restricted (north of Minas Gerais and south of Bahia states, Brazil); (ii) its diverse morphological traits compared with other *Tarenaya* species; (iii) the absence of the synapomorphies found in other *Tarenaya*; and (iv) the previous molecular phylogenetic evidence indicating the possible separation of *T. siliculifera* from *Tarenaya,* the main goal of our work was to characterize and to shed light on the phylogenetic relationship among Brazilian *Tarenaya* species, culminating with the erection of *Neotarenaya siliculifera* gen. nov. *et* comb. nov. under the provisions of the International Code of Nomenclature for algae, fungi, and plants. To achieve this objective, molecular phylogenetic analysis based on three markers ITS (nuclear), *matk* and *ndhF* (chloroplast) as well as morphological analyses were conducted. In addition, we provide an updated morphological description of *Neotarenaya siliculifera* gen. nov. *et* comb. nov., its geographical distribution and its classification in terms of conservation status, according to the criteria adopted by the IUCN (2017).

2. Material and methods

2.1 Specimens, sampling sites, and morphological analysis

Botanical material of Cleomaceae was obtained from different Brazilian states (Table 1). The herbaria BHCB, CESJ, HUEFS, CEPLAC, FLOR, MBM, UPCB and VIC (acronyms according to Thiers, 2019) were consulted, in an effort to locate additional specimens for analysis and to better understand species distributions. Information in *opus principes* was used to determine the specimens. The vegetative and reproductive structures were analyzed with the aid of a stereoscopic microscope and the measurements were performed with pachymeter, ruler, and millimeter paper. A list of the collected specimens as well as data of their origins and the herbarium voucher are also included in Table 1. Commentaries on the conservation status of *Tarenaya siliculifera* (now *Neotarenaya siliculifera* gen. nov. *et* comb. nov.), described according to the standards of the International Union for Conservation of Nature (IUCN), are provided. A map, showing the distribution of *N. siliculifera* is also presented (Figure 1), which was designed using QGIS 2.14.3 'Essen' (Quantum Gis Development Team, 2016). The main morphological characteristics of the species are additionally presented in Figure 2.

2.2 DNA extraction, amplification, and sequencing

Total DNA was extracted from fresh samples (Table 1) as well as from dried-herbarium material (HUEFS 92185), using a modified CTAB method (Doyle et al., 1990). Three genetic regions (molecular markers) were investigated: two from the chloroplast – cpDNA (*ndhF*, *matk*) and, one from nucleus – nDNA (ITS), considered as phylogenetically informative at the inter- and infrageneric levels in the Cleomaceae family (Inda et al., 2008; Patchell et al., 2014). The segments of interest were amplified by means of polymerase chain reaction (PCR), which were implemented with a total volume of 50 μL: 0.5 ρmol of each oligonucleotide primer, 0.2 mM of each dNTP, 2.0 mM MgCl2, 1 × PCR buffer (20 mM Tris-HCl (pH 8.4), 50 mM KCl), and 1.5 U Platinum Taq DNA polymerase (Invitrogen Life Technologies, Carlsbad, CA, USA). Primers used in this study were based on previous studies (Koch et al., 2001; Olmstead et al., 1993; White, 1990) (Supplementary material – Table S1). The thermal cycling conditions were carried out according to Patchell et al., (2014). PCR products were visualized using 1 % (agar – *w*/*v* on TAE buffer) gel electrophoreses and then cleaned with Wizard SV Gel and PCR Clean-

Up System PCR purification columns (Promega). The generated fragments were purified and submitted to DNA sequencing at Mylleus (https://store.myleus.com/). The sequenced fragments were assembled into one contig using the software Sequencer 4.1 (GeneCodes, Ann Arbor, Michigan).

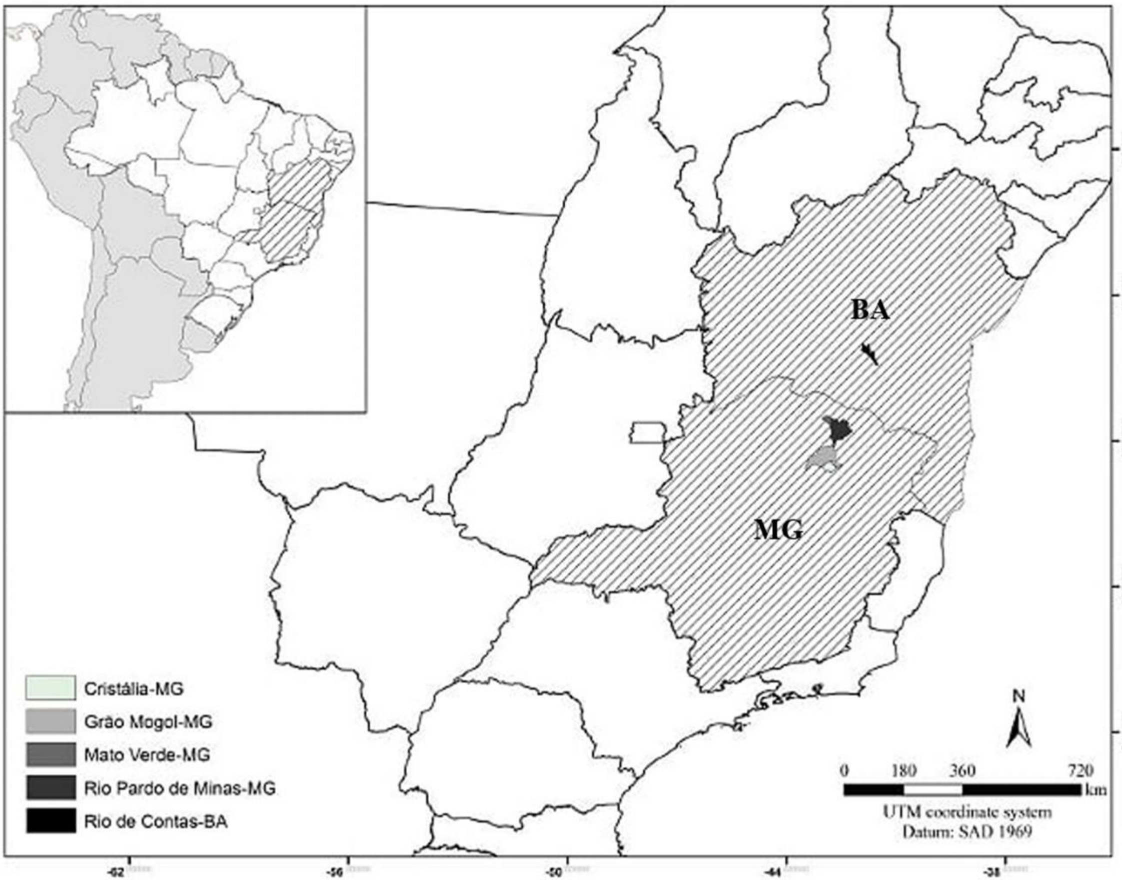


Figure 1: Geographical distribution of *Neotarenaya siliculifera* (Cleomaceae, Brassicales) in both Minas Gerais (MG) and Bahia (BA) states (Brazil) that are highlighted with hatched lines. The cities from which specimens of *N. siliculifera* were collected are located in the "Cadeia do Espinhaço", a mountain range found in Atlantic uplands. The legends refer to cities (city name followed by the state abbreviation).

2.3 Phylogenetic analysis

The nucleotide sequences obtained in this study (partial ITS, *matk* and *ndhF* sequences) and other sequences retrieved from GenBank were used to generate phylogenetic trees (Supplementary material – Table S2). Phylogenetic reconstruction was done on concatenated gene sequences (ITS, *matk* and *ndhF*). The sequences (136) were aligned and trimmed (matrix with a 4,074-bp length) using *CLUSTAL W*. After, phylogenetic inferences were carried out, based on the Maximum Likelihood (ML), Bayesian inference (BI) and

Maximum Parsimony (MP) methods. In addition, phylogenetic trees for the individual molecular markers were also constructed, which were evaluated by BI. Thus, for the ITS nuclear marker, a matrix with 114 taxa and 1,172-bp length was used, whereas for the chloroplast markers *matk* and *ndhF*, 86 taxa and 1,424-bp length as well as 74 taxa and 1,478-bp length were used, respectively.

The ML analysis were carried out using RAxML version 8.2.10 and node support was obtained using the rapid bootstrap algorithm with 1,000 replicates (Stamatakis, 2014). The search for the best-scoring ML tree was performed with the general time-reversible model of DNA sequence evolution with gamma-distributed rate heterogeneity (GTRGAMMA model). The MP analysis was conducted with tree bisection-reconnection (TBR) branch swapping and the initial trees were generated by 10,000 random stepwise addition replicates. The strict consensus trees were obtained and support for each node was assessed by bootstrap analysis with 2,000 replicates through the heuristic search method. The BI analyzes was performed according to the best fit evolutionary model which was selected by the Akaike information criterion (AIC) (Posada and Buckley, 2004) using MrModeltest 2.2 (Nylander, 2004). The Markov chain Monte-Carlo algorithm was executed with four runs with 50 million generations each and sampling every 5000 generations. The first 25% of the trees were discarded as burn-in, the remaining trees were used to construct the majority rule consensus tree, and then the PP for each node was calculated.

3. Results

3.1 Phylogenetic analysis within *Tarenaya*

A comparative analysis among the novel gene sequences and those available in GenBank revealed similarities ranging from 85 to 98.9%. In addition, all novel sequences were related to ITS, *matk* and *ndhF* ones retrieved from Cleomaceae species. Maximum Likelihood (ML), Bayesian inference (BI) and Maximum Parsimony (MP) analyses showed slightly different topologies, however each supported the 15 major clusters found previously (Barrett et al., 2017; Patchell et al., 2014). As summarized in the BI consensus tree, based on the concatenated ITS, *matk* and *ndhF* sequences, Cleomaceae species are distributed in, at least, 22 clusters (Fig. 3). Most of these clusters are monophyletic groups (except *Cleome* s.l.). In general, the phylogenetic grouping of species/genera was congruent with their geographical distribution.

The genera *Peritoma* and *Cleomella*, found in North America, were grouped as a sister cluster, as observed for Latin American groups (*Cleoserrata, Tarenaya, Melidiscus, Andean, Physostemon* and *Dactylaena*), among others. This relationship was also observed for the phylogenetic reconstructions of single-gen molecular markers (Supplementary Material – Figures S3-S5). However, it is important to highlight that the topologies between the concatenated (Fig 3) and the ITS-based tree (Supplementary Material – Figure S3) are more similar, indicating that ITS sequences are most likely indicative for inferences below the family level, as previously observed (Patchell et al., 2014).

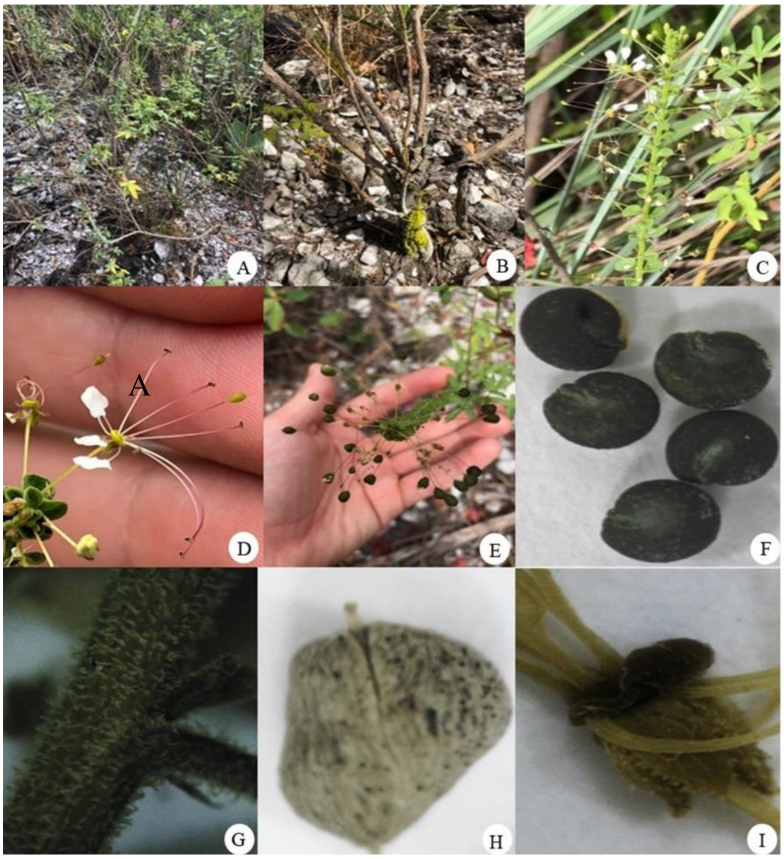


Figure 2: Pictures of *Neotarenaya. siliculifera* showing its natural habit as well as specific morphological traits. A: An individual of *N. siliculifera* in its natural habit (Estadual Park of "Serra Nova" – Parque Estadual Serra Nova, Rio Pardo de Minas municipality, Minas Gerais state, Brazil, 2018). B: Stem. C: Racemous inflorescence. D: Delicate flower. E: Fruit type capsule. F: Seeds. G: Details of the connection of the petiole in the branch with several glandular trichomes and absence of prickly stipules. H: Details of the fruit with apical opening. I: Nectariferous disk at base of corolla.

Table 1: Novel brazilian specimens used in molecular phylogeny analyzes, and their respective collection sites [city-state (acronyms for state abbreviation)] and herbarium vouchers.

Specimens	**City – State**	**Herbarium voucher**
Cleoserrata paludosa	Belém – Pará (PA)	VIC 51788
Gynandropsis gynandra	Natal – Rio Grande do Norte (RN)	VIC 51789
Melidiscus giganteus	Campos Mourão – Paraná (PR)	VIC 51790
Neotarenaya siliculifera	Rio Pardo de Minas – Minas Gerais (MG)	VIC 51801
Neotarenaya siliculifera	Rio de Contas – Bahia (BA)	HUEFS 92185
Tarenaya aculeata	Feira de Santana – Bahia (BA)	VIC 51791
Tarenaya diffusa	Feira de Santana – Bahia (BA)	VIC 51793
Tarenaya hassleriana	Canaã – Minas Gerais (MG)	VIC 51795
Tarenaya hassleriana	Domingos Martins – Espírito Santo (ES)	VIC 51907
Tarenaya hassleriana	Canoinhas – Santa Catarina (SC)	VIC 51905
Tarenaya hassleriana	Joinville – Santa Catarina (SC)	VIC 51802
Tarenaya hassleriana	Piau – Minas Gerais (MG)	VIC 51904
Tarenaya hassleriana	São Miguel do Anta – Minas Gerais (MG)	VIC 51796
Tarenaya hassleriana	Viçosa – Minas Gerais (MG)	VIC 51794
Tarenaya longicarpa	Picos – Piauí (PI)	VIC 51799
Tarenaya microcarpa	Belém – Pará (PA)	VIC 51797
Tarenaya parviflora	Pombal – Paraíba (PB)	VIC 51798
Tarenaya rosea	Colatina – Espírito Santo (ES)	VIC 51800
Tarenaya sp.	Afrânio – Pernambuco (PE)	VIC 51900
Tarenaya sp.	Arcoverde – Pernambuco (PE)	VIC 51903
Tarenaya sp.	Ibimirim – Pernambuco (PE)	VIC 51901
Tarenaya sp.	Lavras – Ceará (CE)	VIC 51902
Tarenaya sp.	Manaus – Amazonas (AM)	VIC 51792
Tarenaya spinosa	Teresina – Piauí (PI)	VIC 51804

Based on the phylogeny of concatenated sequences, the cluster *Tarenaya* s.l., harboring 39 species, and having *Cleoserrata* as a sister group, was divided into two sub-clusters (*Tarenaya* s. s. and *Ser.* Cleorosea), which were strongly supported (ML=100, BI=100) (Fig. 3). Thus, three of the four series of *Tarenaya* (Parviflorae, Aculeatae/Hemiscola and Spinosa) were grouped in the first cluster (*Tarenaya* s.s.), while members of the series Cleorosea, formed by the species *T. rosea* and *T. siliculifera* (*Neotarenaya*) were grouped together (Fig. 3; Supplementary Material – Figures S1 and S2). Accordingly, the series of *Tarenaya*, except Cleorosea, are presented as polyphyletic groups (Supplementary Material – Figures S1 and S2). It is important to mention that regarding the phylogeny inference based on ITS sequences and on concatenated sequences, the genus *Tarenaya* is monophyletic (Fig. 3; Supplementary Material – Figure S3). However, we observed in the molecular phylogeny based on the chloroplastic *matk* sequences that the genus *Tarenaya* is polyphyletic, since a group composed mainly by species isolated from Brazil has grouped into a distinct cluster from *Tarenaya* s.s (Supplementary Material – Figure S4). The marker *matk* is highly conserved and thus not useful at lower taxonomic levels, compared to the concatenated-sequences tree and the ITS-based tree with respect to the *Tarena* s.s., *Neotarenaya* and *Cleoserrata* groups. Finally, in the *ndhF*-based tree (Supplementary Material – Figure S5), the topology is rather similar to that observed in the concatenated one, although without the *Cleoserrata* group, which is not represented in this phylogenetic reconstruction.

Considering the phylogenetic reconstructions based on concatenated sequences (Fig. 3) as well as in single genes (Supplementary Material – Figures S3-S5), the sequences of the proposed new genus, *Neotarenaya siliculifera*, were closely related to *Tarenaya* s. s. and *T. rosea*. This fact was more evident for the concatenated tree and for the ITS-based, in which *N. siliculifera* and *T. rosea* were found as sister groups (Fig. 3; Supplementary Material – Figure S3). Despite this phylogenetic relatedness, the sequences from these species were not placed in the same cluster and morphologically these two morphospecies are highly different from each other (Table 2), thus sustaining the separation of the genera *Tarenaya* and the proposal of *Neotarenaya* as a new generic entity. Accordingly, similarly to other genera recently created within the family Cleomaceae (*e.g. Stylodocleome* and *Thulinella* – Roalson et al., 2015 – and *Areocleome* – Barrett et al., 2017), *Neotarenaya*, thus far, is a monotypic genus.

3.2 Morphology

Neotarenaya siliculifera presents, when adult, leaves only at the apex of the branches, in addition it has a twisted stem that is thick at the base, and branched in the median region (Fig. 2A-B). The morphological vegetative characters (*e.g.* leaves) and reproductive (*e.g.* flowers and fruits) are tiny (a few millimeters long). The species has a high density of glandular trichomes on the leaves and inflorescences, giving them a whitish appearance. In addition, as aforementioned, *N. siliculifera* does not have spiny stipules (Fig. 2G), the fruit is small containing only four seeds (Fig. 2H) and the seeds lack the aryl observed in *Tarenaya* (Fig. 2F).

In addition, the novel genus *Neotarenaya* diverges from *Tarenaya* since it does not contain the morphological synapomorphies found in *Tarenaya*, mainly the presence of a pair of spines at the base of the petiole. More importantly, the specimens belonging to the novel genus *Neotarenaya* display the following synapomorphic traits: tiny leaves and flowers, a few millimeters long; fruits of capsule type, with only four seeds without aryl; the specimens also show a high density of glandular trichome on the leaves, without spiny stipules, and inflorescences giving a whitish appearance. However, as observed in the phylogenetic positioning, which indicates an evolutionary proximity between members of *Neotarenaya* and *Tarenaya*, which are sister clusters, some morphological characters are shared by these genera, such as: unguiculata petals (Fig. 2D), a nectariferous disc (Fig. 2I) and bracteada at the base of the corolla and inflorescence (Fig. 2C).

3.3 Taxonomic treatment

Neotarenaya Parma DF., Vaz MGMV. & Nunes-Nesi A., *gen. nov.* – TYPE: *Neotarenaya siliculifera* (Eichler) Parma DF., Vaz MGMV. & Nunes-Nesi A. (*Cleome siliculifera* Eichler, Fl. Bras. 13(1): 260. 1865.) [BRAZIL. Minas Gerais] Prov. Minarum: Habitat prope Cocaes et in Serra dos Pinheiros, [1839], J. Pohl s.n. (lectotype, image!).

The following set of characters distinguishes this novel taxon from all other species placed in Tarenaya, suggesting its description as a new generic entity in the Cleomaceae family: The leaves and flowers are tiny, with a few millimeters long. Fruits type capsule with only four seeds without aryl. The species shows a lot of glandular trichome in the leaves, without spiny stipules, and inflorescences giving a whitish appearance.

Habit: shrub 0.7-1.84 m height, with glandular trichomes. **Leaves:** petioles 0.79-1.67 cm long, glandular-pubescent; leaflets 3-5, dark green; sessile lamina, linear-lanceolate, glandular-pubescent, ciliated margins, camphorma venation; terminal leaflet: obtuse apex and softened base, 0.67-1.46 cm long, 0.28-0.29 cm wide; pairs of lateral and basal leaflets: obtuse apex and softened base, the first pair 0.28-0.47 cm long, 0.14-0.24 cm long, the second pair 0.25-0.31 cm long, 0.14-0.22 cm long. **Inflorescence:** racemous; peduncle brown, glandulous-pubescent, 2.5-3.4 cm length; bracts simple with petiole 0.5-0.7 mm long, glandular-pubescent, dark green, ovate-elliptic, glandular-pubescent, ciliated margin, obtuse apex, truncated base, 2.20-3.26 mm long, 1.77-2.78 mm wide, camphoroproma venation, pedicel lilac, glandulous-pubescent, 0.71-1 mm length; **Flower:** sepals lanceolate, reflexes after anthesis, 0.43-0.56 mm long, 0.63-0.81 mm wide, light green, glandular-pubescent on the abaxial face, glabrous adaxial

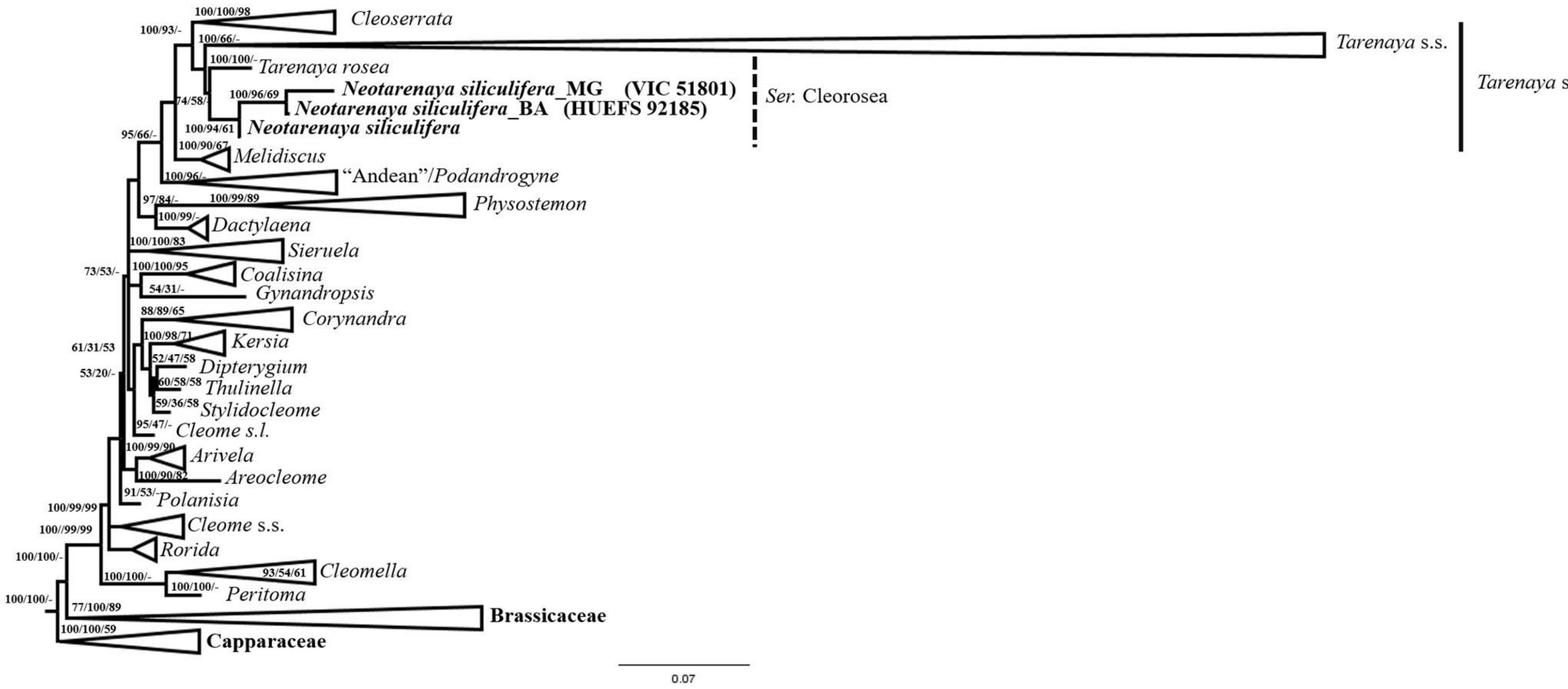


Figure 3: Phylogenetic relationships between and within Capparaceae, Brassicaceae and Cleomaceae by Bayesian inference of concatenated sequences of nuclear ITS and chloroplastic *matK* and *ndhF*. Numbers at nodes reflect Posterior Probability (PP) from Bayesian inference, Bootstrap values from Maximum Likelihood (ML) and Bootstrap values from Maximum Parsimony (MP): PP/ML/MP. Bar: 0.07 substitutions per nucleotide.

face; petals unguiculata-linear, nail 2-5 mm length, orbicular to elliptic, rounded apex, 1.41-2.0 mm long, 0.56-0.78 mm long; stamens 6, purple; fillets 5.5-7.23 mm long, glabrous; anther 0.55-0.74 mm length; disc nectariferous conical to globose, 0.27-0.4 mm length; gauge 5.50-7.23 mm; ovary green, 0.55-0.72 mm length; sessile stigma, discoid, 0.08-0.09 mm, ovoid. **Fruit:** glabrous, 2.69-3.09 mm long, 1.31-1.38 mm long, 4 seeds. **Seeds**: ovoid orbicular, 1.31-1.35 mm long, black and smooth.

Phenology: Flowering occurs between the months of October (SPF 35188) and April (SPF 22827; VIC 51801).

Geographic distribution and habitat: *Neotarenaya siliculifera* is known from Minas Gerais and Bahia states (Fig. 1), where it occurs in a transitional zone between the upper montane forest and high altitude grassland habitat, on rocky outcrops or in rocky soils.

Etymology: "Neo" comes from the Greek neos, which means "new." Therefore, *Neotarenaya* represents a new genus coming from *Tarenaya*.

Conservation status: We applied the IUCN (2017) criteria and propose an IUCN red list category of vulnerable (VU=D2) given the low number of locations where this species was found (equal to five). There are few records for this species and most of them are old (1839, 1977, 1981, 1985, 1988, 2004, 2006 and 2012). In addition, only one site of occurrence of the species corresponds to an area of environmental preservation, being the others close to cities.

Additional specimens examined: BRAZIL: Bahia: Rio de Contas, in the base of “Serra do Tombador”, 974 m, lat. -13.627778 and long. -41.8 WGS84, 10 march 2004, Harley, R.M. 55098 (HUEFS 92185); Waterfall “Fraga do Rio Brumado”, outskirts of the city, 900 m, lat. -13,58 and long. -41.8 WGS84, 24 November 1988, Harley, R.M. 26987 (NY 485327); Vila do Rio de Contas, on side road to Livramento, leading to the Rio Brumado, 980 m, 28 march 1977, Harley, R.M. 20135 (UEC 37874). Minas Gerais: Grão Mogol, Vale do Rio Itacambiruçu (Itacambiruçu River Valey), between the “Jambeiro” farm and the road to Cristália, 700m, lat. -16.59 and long. -42.88 WGS84, 26 February 1986, Chukr, N.S. et al. 9698 CFCR (UEC 45549); Mato Verde, about 14 km to Monte Azul, coming from Montezuma, 924m, lat. -15.39 and long. -42.78 WGS84, 11 January 2006, Santos, A.K.A. 626 (HUEFS 104737); “Grão-Mogol-Cristália” Road, about 2 km from the junction with the “Boqueirão-Grão-Mogol” road, 650 m, lat. -16.59 and long. -42.75 WGS84, 14 February 2003, França, F. 4352 (HUEFS 70286); Along the bridge over the Rio das Mortes (das Mortes river), 690 m, lat. -16.60 and long. -42.90 WGS84, 24 january 2002, Schutz et al.1 1359 (UEC 123752); Rio Pardo de Minas, Serra Nova State Park, Sand trail towards Biu biu stream, lat: -15.663611 and long: -42.770556

WGS84, 26 march 2012, Rocha, MJR et al. 582 (BHCB 154742), 21 April 2018, Parma, DFP & Vinicíus-Silva R. 72 (VIC 51801).

Notes: *Neotarenaya siliculifera* was classified as *Tarenaya* by Luciano et al., (2018). The authors lectotypify *C. siliculifera* with BR 698586.

4. Discussion

4.1 Phylogenetic relationships within Cleomaceae

Several systematic and taxonomic questions remains within the family Cleomaceae, including, for instance, relationships and nomenclature within the genus *Tarenaya*, (*e.g.* Andean cluster) (*e.g.* Barrett et al., 2017; Roalson et al., 2015, Thulin and Roalson, 2017). Nevertheless, there remains much to be done, since various inaccuracies/conflicts are still existing, such as infra-generic delimitations, as well as monotypic genera not covered by molecular studies (*e.g. Haptocarpum* Ule and *Puccionia* Chiov.). Furthermore, certain morphologically distinct genera, comprised by one or few species are embedded within larger genera, leading to incongruences and to the establishment of non-monophyletic groups. This pattern suggests that most, if not all, of these small genera need to be sampled more intensively, enabling a better understanding of phylogenetic relationships and generic delimitations within these groups (Hall, 2008).

Some large genera have been divided into smaller ones, such as *Areocleome, Corynandra, Gilgella, Dipterygium, Kersia, Puccionia, Stylidocleome* and *Thulinella*. In this way, it becomes easier to recognize/identify each cluster, which in addition to presenting well-defined morphological characteristics also have strong phylogenetic support. The genus *Tarenaya*, which is relatively large, could be separated into small genera, rather than divided into series. However, for such a modification a more comprehensive study is clearly required, with a more complete sampling of the genus.

4.2 Phylogenetic relationships within *Tarenaya*

The genus *Tarenaya* as currently recognized contains 40 species, spread into four series (Luciano et al., 2018a). Accordingly, the current classification of this genus requires revision since, as observed by Feodorova et al., (2010) and also here, molecular, morphological and ecological data strongly support internal groups within *Tarenaya*.

Given the geographic distribution of *Tarenaya*, two main groups are formed: the cluster Spinosae (*Tarenaya* s.s) with a wide occurrence in the tropics, whereas Hemiscola and Parviflorae have narrower distributions in the Neotropics. The species from cluster Spinosae are larger sized (e.g. shrubs), showy flowers and occur in all types of environments. In contrast, neotropic species are generally endemic to one determined region, and do not occur throughout the Neotropics (*e.g.* serie Cleorosea – *Neotarenaya*). In this sense, *Neotarenaya* is circumscribed to the north of Minas Gerais and south of Bahia states (Brazil), in an area that corresponds to the "Cadeia do Espinhaço", a mountain range found in Atlantic uplands, and some of the refugia zones of the Atlantic forest. The "Cadeia do Espinhaço" is a mountainous area considered of great ecological importance, since it has many endemic flora species as well as due to its great wealth of habitats. In the past, humid forests would retract during periods of maximum ice (forming refuges) and expand during warmer periods (interglacial periods), while open vegetation areas behave in opposite ways (see theory of refuges in Carnaval et al., 2009). Thus, species dependent on forest environments would have followed these cycles of retraction and expansion of forests during the Pleistocene, leading populations associated with different refuges to diversify in an allopatric manner, if the time of isolation was sufficient to accumulate differences between them (Carnaval et al., 2009; Carnaval and Moritz, 2008).

Table 2: Morphological comparison of *Tarenaya rosea* and *Neotarenaya siliculifera*

Character	*T. rosea*	*N. siliculifera*
Habit	herb	shrub
Modified stipules in spines	present	absent
Terminal leaf blade (length)	1.65-6.8 cm	0.67-1.46 cm
Floral bract (length)	2-11 mm	2.20-3.26 mm
Fruit (type)	silica	capsule
Ovary	cylindrical	ovoid
Seed surface	ribbed	smooth
Seeds per fruit	50	4

4.3 *Neotarenaya siliculifera*

The genus *Tarenaya*, native from America, is recognized as monophyletic (based on molecular marker ITS), and comprises about 40 species (Luciano et al., 2018a; Patchell et al., 2014). *Tarenaya* taxa share distinctive morphological traits such as a closed seed cleft covered by a thin membrane, spinolose pollen grains (Inda et al., 2008) and a pair of spines at the base and along the petiole (Luciano et al., 2018a); characteristics which are not shared by the novel genus, *Neotarenaya.* Our molecular phylogenetic results demonstrated that the *Tarenaya* cluster would be one of the most derived in the family (Fig. 3). Notwithstanding, *Neotarenaya* presents fruit of the type a capsule, with only four seeds, which is a morphological characteristic considered as less derived, compared to the siliqua fruit observed in the other species of the family.

5. Conclusion

The new genus and the new species *Neotarenaya siliculifera* gen. nov. *et* comb. nov., which was erected considering its phylogenetic placement (based on ITS, *matk* and *ndhF*) and morphological traits, has a restricted geographical distribution. This novel species is circumscribed to the north of Minas Gerais and south of Bahia states (Brazil), in an area that corresponds to the "Cadeia do Espinhaço" and some of the refugia zones of the Atlantic forest. The "Cadeia do Espinhaço" is a mountainous area of great ecological importance that presents many species of endemic flora, due to its great wealth of habitats. Accordingly, this description highlights the need for more floristic surveys in this underexplored region. More importantly, the sampled material must be deposited in public herbaria and, if possible, be cultivated, allowing future studies not only on reproduction and pollination but also on metabolism and physiology.

Acknowledgments

This work was supported by funding from the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) (grant number 424024/2018-7), and the Foundation for Research Assistance of the Minas Gerais State (Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG) (grant number APQ-00528-18). Research fellowships granted by CNPq to ANN and WLA as well as the scholarship granted by the Coordination of Improvement of Higher Education Personnel (CAPES) to DFP are also gratefully acknowledged. MGMVV was supported by scholarships from CAPES (PNPD-1638006). APMW acknowledges funding by the Deutsche

Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2048/1 – project ID 390686111.

References

Barrett, R.L., Roalson, E.H., Ottewell, K., Byrne, M., Govindwar, S.P., Yadav, S.R., Tamboli, A.S., Gholave, A.R., 2017. Resolving generic boundaries in Indian-Australasian Cleomaceae: Circumscription of *Areocleome*, *Arivela*, and *Corynandra* as distinct genera. Syst. Bot. 42, 694–708. https://doi.org/10.1600/036364417X696401

Bayat, S., Schranz, M.E., Roalson, E.H., Hall, J.C., 2018. Lessons from Cleomaceae, the Sister of Crucifers. Trends Plant Sci. 23, 808–821. https://doi.org/10.1016/j.tplants.2018.06.010

Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., Moritz, C., 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science (80-). 323, 785–789. https://doi.org/10.1126/science.1166955

Carnaval, A.C., Moritz, C., 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. J. Biogeogr. 35, 1187–1201. https://doi.org/10.1111/j.1365-2699.2007.01870.x

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Feodorova, T.A., Voznesenskaya, E. V, Edwards, G.E., Roalson, E.H., 2010. Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). Syst. Bot. 35, 811–826. https://doi.org/10.1600/036364410X539880

Flora of Brazil, 2020. Flora of Brazil 2020 - Algas, Fungos e Plantas [WWW Document]. URL http://reflora.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/BemVindoConsultaPublicaConsultar.do?invalidatePageControlCounter=1&idsFilhosAlgas=%5B2%5D&idsFilhosFungos=%5B1%2C11%2C10%5D&lingua=&grupo=5&genero=Tarenaya&especie=hassleriana&autor=&nomeVer

Guidelines for Using the IUCN Red List Categories and Criteria The Iucn Red List of Threatened Species^{TM}, 2017.

Hall, J.C., 2008. Systematics of Capparaceae and Cleomaceae: an evaluation of the generic delimitations of Capparis and Cleome using plastid DNA sequence data. Botany. 696, 682–696. https://doi.org/10.1139/B08-026

Hall, J.C., Sytsma, K.J., Iltis, H.H., 2002. Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. Amer. J. Bot. 89, 1826–1842. https://doi.org/10.3732/ajb.89.11.1826

Iltis, H.H., Hall, J.C., Cochrane, T.S., Sytsma, K.J., 2011. Studies in the Cleomaceae I. On the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae [1]. Ann. Missouri Bot. Gard. 98, 28–36. https://doi.org/10.3417/2007017

Inda, L.A., Torrecilla, Æ.P., Catala, Æ.P., Ruiz-zapata, T., 2008. Phylogeny of *Cleome* L. and its close relatives *Podandrogyne* Ducke and *Polanisia* Raf. (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. Pl. Syst. Evol. 274, 111–126. https://doi.org/10.1007/s00606-008-0026-y

Koch, M., Haubold, B., Mitchell-Olds, T., 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. Am. J. Bot. 88, 534–44.

Luciano, R., Neto, S., Thomas, W.W., Regina, M., Barbosa, D.V., Roalson, E.H., 2018. New combinations and taxonomic notes for *Tarenaya* (Cleomaceae). Acta Bot. Bras. 32, 540–545. https://doi.org/10.1590/0102-33062017abb0417

Nylander, 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Olmstead, R.G., Jennifer, A., Wolfe, K.H., Biology, E.P.O., 1993. Update section Ninety extra nucleotide in ndhF gene of tobacco chloroplast DNA: a summary of revisions to the 1986 genome sequence. Pl. Molec. Biol. 22, 1191–1193.

Patchell, M.J., Roalson, E.H., Hall, J.C., 2014. Resolved phylogeny of Cleomaceae based on all three genomes. Taxon. 63, 315–328.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808. https://doi.org/10.1080/10635150490522304

Roalson, E.H., Hall, J.C., Ii, J.P.R., Cardinal-mcteague, W.M., 2015. A revision of generic boundaries and nomenclature in the North American cleomoid clade (Cleomaceae). Phytotaxa. 205, 129–144.

Roalson, E.H., Hall, J.C., 2017. New Generic Concepts for African Cleomaceae. Syst. Bot. 42, 925–942. https://doi.org/10.1600/036364417X696393

Sanchez-Acebo, L., 2005. A phylogenetic study of the New World *Cleome* (Brassicaceae, Cleomoideae). Ann. Missouri Bot. Gard. 92, 179–201.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stevens, P.F., 2001 Angiosperm Phylogeny Website. Version 14, July 2017 [and more or less continuously updated since] [WWW Document]. URL http://www.mobot.org/mobot/research/apweb/orders/brassicalesweb.htm (accessed 2.27.19).

Thiers, B., 2019. Index Herbariorum [WWW Document]. URL http://sweetgum.nybg.org/science/ih/ (accessed 3.7.19).

Thulin, M., Roalson, E.H., 2017. Resurrection of the Genus *Rorida* (Cleomaceae), a Distinctive Old World Segregate of *Cleome*. Syst. Bot. 42, 569–577. https://doi.org/10.1600/036364417X695989

White, T.J., Bruns, T.D., Lee, S.B., Taylor J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a Guide to Methods and applications. 38, 315–322.

CHAPTER 3

New insights into the evolution of C_4 photosynthesis offered by the Tarenaya cluster of Cleomaceae

Daniele F. Parma[1], Marcelo G. M. V. Vaz[1], Priscilla Falquetto[1], Jéssica C. Silva[2], Wellington R. Clarindo[2], Philipp Westhoff[3], Robin Van Velzen[4], Urte Schlüter[5], Wagner L. Araújo[1], M. Eric Schranz[4], Andreas P. M. Weber[5], & Adriano Nunes-Nesi[1]*

[1] Departamento de Biologia Vegetal, Universidade Federal de Viçosa, 36570-900 Viçosa, Minas Gerais, Brazil

[2] Departamento de Biologia Geral, Universidade Federal de Viçosa, 36570-900 Viçosa, Minas Gerais, Brazil

[3] Plant Metabolism and Metabolomics Laboratory, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University, Düsseldorf, Germany

[4] Biosystematics Group, Wageningen University, Wageningen, The Netherlands

[5] Institute of Plant Biochemistry, Cluster of Excellence on Plant Science (CEPLAS), Heinrich Heine University, Düsseldorf, Germany

*Correspondence:
nunesnesi@ufv.br

ABSTRACT

Cleomaceae is closely related to Brassicaceae and includes C_3, C_3-C_4, and C_4 species. Thus, this family represents an interesting system for studying the evolution of the carbon-concentrating mechanism. However, inadequate genetic information on Cleomaceae limits their research applications. We characterized 22 Cleomaceae accessions [3 genera (*Cleoserrata*, *Gynandropsis*, and *Tarenaya*) and 11 species] in terms of genome size; molecular phylogeny; and anatomical, biochemical, and photosynthetic traits. We clustered the species into seven groups based on genome size. Interestingly, despite clear differences in genome size (2C, ranging from 0.55 to 1.3 pg) in *Tarenaya*, this variation was not consistent with phylogenetic grouping based on the internal transcribed spacer markers, suggesting the occurrence of multiple polyploidy events within this genus. Moreover, only *G. gynandra*, which possesses a large nuclear genome, exhibited the C_4 metabolism. Among the C_3 species, we observed intra- and interspecific variation in nuclear genome size as well as biochemical, physiological, and anatomical traits. Furthermore, the C_3 species increased venation density and bundle sheath cell size, compared with the C_4 species, which likely predisposed the former lineages to C_4 photosynthesis. Our findings demonstrate the potential of *Tarenaya* in offering novel insights into the evolution of C_4 photosynthesis.

Keywords: Cleomaceae, *Cleoserrata*, genome size, *Gynandropsis*, intermediate photosynthetic mechanism, polyploidy, *Tarenaya*

INTRODUCTION

C_4 photosynthesis is a complex trait with a high degree of natural variation. It has evolved independently over 60 times in 19 different botanical families, involving anatomical and biochemical changes relative to the ancestral C_3 state (Sage, 2004; Sage *et al.*, 2011, 2014). Accordingly, the different origins of this photosynthetic mechanism have been tracked by means of molecular markers (Patchell *et al.*, 2014; Bayat *et al.*, 2018), anatomical traits (e.g., different types of Kranz anatomy; Koteyeva *et al.*, 2011), and biochemical patterns (three types of decarboxylation enzymes; Hatch 1988; Wang *et al.*, 2014). In addition, the transition from C_3 to C_4 photosynthesis did not likely proceed through a single step but rather via a series of transitory stages (Sage 2004; Gowik and Westhoff 2011; Schuler, Mantegazza, and Weber 2016). These include the development of larger bundle sheath cells (BSCs), increase in leaf venation density (VD), restriction of glycine decarboxylase to BSCs, establishment of a photorespiratory carbon pump, enhancement of phosphoenolpyruvate carboxylase activity, establishment of the C_4 cycle, and optimization of the C_4 syndrome (Gowik and Westhoff, 2011; Sage, 2004). In this way, plants exhibiting C_3-C_4 intermediate characteristics bridge the evolutionary gap between the C_3 and C_4 species (Bräutigam and Gowik 2016; Heckmann *et al.*, 2013; Mallmann *et al.*, 2014). Furthermore, phylogenetic studies have shown that many C_3-C_4 plants are closely related to C_4 plants (McKown *et al.*, 2005; Marshall *et al.*, 2007; Christin *et al.*, 2011a; Kadereit and Freitag, 2011; Patchell *et al.*, 2014).

Cleomaceae species have been used as models to study the evolution of C_4 photosynthesis, as this family comprises representatives of the three photosynthetic types (C_3, C_3-C_4 intermediates, and C_4) (Brown *et al.*, 2005; Marshall *et al.*, 2007). Additionally, Cleomaceae is closely related to Brassicaceae (sister families), which includes the C_3 model plant *Arabidopsis thaliana*. Accordingly, the genomes of species in these families show significant synteny (Schranz and Mitchell-Olds, 2006), and this close relationship has facilitated the identification of Cleomaceae orthologs in *Arabidopsis thaliana.* Cleomaceae comprises nearly 200 species (Bayat *et al.*, 2018; Stevens, 2001), three of which have been characterized as C_4 species (*Gynandropsis gynandra, Coalisina angustifolia, and Areocleome oxalidea*), one as a C_3-C_4 intermediate species (*Coalisina paradoxa*), and the rest as C_3 species (Marshall *et al.*, 2007; Voznesenskaya *et al.*, 2007; Koteyeva *et al.*, 2011; Bayat *et al.*, 2018). The three Cleomaceae members characterized as C_4 species have independent origins, as supported by phylogenetic (Patchell *et al.*, 2014), anatomical (Koteyeva *et al.*, 2011, 2014), and physiological evidence (Voznesenskaya *et al.*, 2018). Therefore, this family is promising for obtaining a deeper understanding of the evolution of C_4 photosynthesis; however, only 15 Cleomaceae species have been thoroughly characterized as yet (Marshall *et al.*, 2007; Voznesenskaya *et al.*, 2007; Koteyeva *et al.*, 2011; Bayat *et al.*, 2018).

Furthermore, some C_3 species of Cleomaceae deviate from the canonical C_3 state, exhibiting variable traits associated with the C_4 mechanism, such as increase in leaf VD, enlargement of bundle sheath cells, proliferation of both mitochondria and chloroplasts, and accumulation of transcripts and proteins required for C_4 photosynthesis, similar to C_3-C_4 intermediate and C_4 species (Marshall *et al.*, 2007). These deviations indicate that the species may have a higher propensity for the evolution of C_4 photosynthesis (Marshall *et al.,* 2007), facilitating the acquisition of novel traits (Sage, 2001; Marshall *et al.*, 2007; Mckown and Dengler, 2007; Christin *et al.*, 2011, 2012, 2013), as observed in grasses (Christin *et al.*, 2013a,b,c)(Christin et al., 2013b). Of note, Cleomaceae species exhibiting C_3 (*Tarenaya hassleriana*) and C_4 (*G. gynandra*) photosynthetic metabolism share the copy number of various genes essential for C_4 photosynthesis. However, expression analysis of C_4 photosynthetic orthologs have demonstrated that the regulation of transcript abundance is much less stringent in *T. hassleriana* than in *G. gynandra* (Bergh *et al.,* 2014). Thus, compared with Brassicaceae species, Cleomaceae species are expected to be predisposed to the evolution of C_4 photosynthesis, as already demonstrated for *G. gynandra* (Huang *el al.*, 2021), given the evolutionary history of the group, including the occurrence of ancient whole genome duplications (Th-α) and paleopolyploidy events (Bergh *et al.*, 2014).

Cleomaceae species with C_4 photosynthesis have independent origins. The *Coalisina* group includes C_3, C_4, and C_3-C_4 intermediate species. The two other groups, namely *Gynandropsis* and *Areocleome*, do not include C_3 or C_3-C_4 intermediate species in their respective clusters, since they are monotypic genera (Patchell *et al.,* 2014; Bayat *et al.,* 2018). In this context, additional species of the family need to be characterized, as species with different photosynthetic metabolism within the same group may facilitate the understanding of the acquisition/evolution of the C_4 metabolism (Bayat *et al.*, 2018). Although *G. gynandra*—the best-studied C_4 species of Cleomaceae—does not have a close C_3 or C_3-C_4 intermediate relative, studies with different accessions of the species have proven the presence of variations in the C_4 metabolism (Reeves *et al.*, 2018). Establishment of a mapping population enables molecular marker trait association using methods such as quantitative trait locus mapping and genome-wide association studies (Reeves *et al.*, 2018).

Despite the greatest species diversity (Stevens 2001), few Cleomaceae species have been sampled for physiological traits to date. Brazil, for instance, is home to 34 Cleomaceae species (Flora of Brazil 2020), most of which are endemic, but no species or accession has been physiologically characterized thus far. However, previous carbon isotope composition analysis has demonstrated the potential of this family to comprise species with the C_4 metabolism (*e.g.*, *T. siliculifera*, a Brazilian endemic species) (Voznesenskaya *et al.,* 2007).

Given the above, knowing that Brazil has a great diversity in species of Cleomaceae and presents arid regions that would be favorable for the development of species with the C_4 photosynthetic mechanism; we tested the hypothesis that "Brazil has species of Cleomaceae with the C_4 photosynthetic mechanism". To this end, the present study aimed at characterizing and identifying the type of photosynthetic metabolism in Brazilian Cleomaceae species and, thus, gaining insights into the evolution of C_4 photosynthesis in this family. We investigated the phenotypic variability in terms of physiological (gas exchange), anatomical (diaphanization and cross section), and metabolic (sugars, starch, nitrogen compounds, and organic acids) traits in 22 phenotypically diverse accessions of Cleomaceae [distributed in 3 genera (*Cleoserrata*, *Gynandropsis*, and *Tarenaya*) and 11 species, including seven *T. hassleriana* accessions and six *T. longicarpa* accessions]. In addition, we measured the genome size of the studied species and created a molecular phylogeny based on the internal transcribed spacer (ITS) nuclear marker. The results are discussed in the context of the C_4 evolutionary aspects of Cleomaceae species.

MATERIAL AND METHODS

Taxon sampling

For a subset of 22 Cleomaceae species selected for the present study, comprising herbaceous and shrub species, seeds were collected *in situ* from 20 locations distributed in 11 Brazilian states (Table 1). Seven accessions were classified as *T. hassleriana* (shrub species), collected from different environments [locations: Viçosa, São Miguel do Anta, Canaã, and Piau in the state of Minas Gerais (MG); Domingos Martins in the state of Espírito Santo (ES); and Joinville and Canoinhas in the state of Santa Catarina (SC)]. Six accessions were classified as *T. longicarpa* (shrub species), collected from different environments [locations: Manaus in the state of Amazonas (AM); Ibimirim, Afrânio and Arcoverde in the state of Pernambuco (PE); Picos in the state of Piauí (PI); and Lavras in the state of Ceará (CE)]. The other morphospecies were classified as *C. paludosa* (shrub species) and *T. microcarpa* (herb species) collected from Belém in the state of Pará (PA); *T. aculeata* (herb species) and *T. diffusa* (herb species) collected from Feira de Santana in the state of Bahia (BA); *T. rosea* (shrub species) collected from Colatina (ES); *T. spinosa* (shrub species) collected from Teresina (PI); *T. parviflora* (shrub species) collected from Pombal in the state of Paraíba (B); *T. siliculifera* (shrub species) collected from Rio Pardo (MG); and *G. gynandra* (herb species) collected from Mossoró in the state of Rio Grande do Norte (RN). The exsiccates are deposited in the VIC herbarium of the Universidade Federal de Viçosa (UFV).

The studied genera can be easily distinguished from one another based on their morphological traits. *Cleoserrata* species are herbs with a glabrous surface, bearing leaves with three leaflets but without stipules, ebracteate inflorescences, and large smooth black seeds. *Gynandropsis* species are annual herbs with a puberulent-glandular surface, bearing leaves with three to five leaflets, ebracteate inflorescences, and small smooth black seeds. Meanwhile, *Tarenaya* species are herbs or shrubs with a puberulent-glandular or pubescent surface, bearing three to seven foliolate leaves with epinescent stipules, bracteate inflorescences, and seeds with longitudinal and transverse streaks that are not very well developed (Stevens 2001; Flora of Brazil 2021).

Furthermore, the different *Tarenaya* species in Brazil can be easily distinguished from one another. *T. aculeata* bears tree foliolate leaves, small white petals, and slightly moniliform capsules; *T. diffusa* bears three to five foliolate leaves, small white petals, and fusiform capsules; *T. microcarpa* bears three to five foliolate leaves, glandular-puberulent indumentum, and small purplish petals; *T. hassleriana* bears armed petioles, glandular-puberulent indumentum, spines on the veins, and pink flowers; *T. parviflora* bears glabrescent indumentum, three to seven foliolate leaves, armed petioles, and white to deep pink petals; *T. longicarpa* bears glandular-puberulent indumentum, completely white petals or petals white at base and purple at the apex (depending on accession), and long capsules, with the fruit being two to three times longer than the gynophore (which differs from *T. spinosa*); *T. spinosa* bears glandular-puberulent indumentum and small white petals (compared with *T. longicarpa*); *T. rosea* bears glandular-puberulent indumentum and petals white at base and pink at apex; and *T. siliculifera* bears glandular-puberulent indumentum without stipular spines, small white petals, obovoid capsules (almost all species of the genus have cylindrical fruits). The different accessions of *T. hassleriana* and *T. longicarpa* could also be distinguished from one another based on height, hairiness, spine density and disposition, and inflorescence and flower size.

The collected seeds were germinated in 5 L plastic containers, filled with a commercial substrate supplemented with 14 g of NPK (4:14:8) per pot [0.56 g of N, 0.86 g of P (P_2O_5), and 0.93 g K (K_2O)]. During spring and summer, five plants of each accession were grown in a greenhouse under semi-controlled conditions (maximum photosynthetically active radiation of 1,500 μmol photons m^{-2} s^{-1} and temperature of 30 ± 2°C) and were irrigated daily. Under these conditions, the plants were grown for 2 (herbaceous species) to 5 months (shrubby species) until physiological characterization and harvesting of leaf material for anatomical and biochemical analyses.

DNA extraction, amplification, and sequencing

Total DNA was extracted from 22 fresh samples using a modified CTAB method (Doyle *et al.*, 1990). For phylogenetic analyses, the ITS regions were sequenced using the universal primers ITS 1 and ITS 4, as previously described (White *et al.*, 1990). The fragments of interest were amplified using polymerase chain reaction (PCR). PCR products were visualized using 1% agarose gel electrophoresis, isolated, and cleaned with the Wizard SV Gel and PCR Clean-Up System PCR purification columns (Promega). The generated fragments were purified and sequenced (Myleus; Belo Horizonte, Brazil). The sequence electropherograms were visually assessed, edited, base-called, assembled, and manually aligned using Sequencer 4.1 (GeneCodes, Ann Arbor, Michigan).

We performed a comprehensive phylogenetic analysis by adding the ITS sequences recovered from the National Center for Biotechnology Information. Based on these sequences, we generated a matrix with 135 taxa of Cleomaceae (648 bp in length). The results were interpretated using Bayesian inference (BI). BI analyses were performed in MrBayes using the best-fit evolutionary model, selected according to the Akaike information criterion (Posada and Buckley, 2004), with MrModeltest 2.2 (Nylander, 2004). The Markov chain Monte Carlo algorithm was executed using four runs with 50 million generations each and sampling every 5,000 generations. The first 25% trees were discarded as burn-in. The remaining trees were used to construct the majority rule consensus tree, and the posterior probability (PP) for each node was calculated.

Flow cytometry

Leaf fragments (~2 cm^2) from each Cleomaceae accession (sample) and *Solanum lycopersicum* (internal standard, 2C = 2.00 pg; Praça-Fontes *et al.*, 2011) were co-chopped in 0.5 mL of OTTO-I nuclear extraction buffer (Otto, 1990) supplemented with 2 mM dithiothreitol and 50 µg mL^{-1} RNAse (Praça-Fontes *et al.*, 2011). Then, 0.5 mL of the same buffer was added, and the suspensions were filtered through a 30 µm nylon mesh, placed into a microtube, and centrifuged at 100 ×g for 5 min. The precipitate was resuspended in 100 µL of OTTO-I buffer and incubated for 10 min. The suspensions were stained with 1.5 mL of OTTO-I:OTTO-II (1:2) buffer supplemented with 2 mM dithiothreitol, 75 µg mL^{-1} propidium iodide, and 50 µg mL^{-1} RNAse (Praça-Fontes *et al.*, 2011). The suspensions were incubated in the dark for 30 min, filtered through a 20 µm nylon mesh, and analyzed using a flow cytometer (BD Accuri C6, Accuri Cytometers, Belgium) equipped with a 488 nm laser source to promote propidium iodide emission at FL2 (615 – 670 nm) and FL3 (> 670 nm). Fluorescence peaks of the G_0/G_1 nuclei of each Cleomaceae accession and *S. lycopersicum* were analyzed based on histograms using BD Accuri™ C6. The G_0/G_1 nuclear peaks exhibiting a coefficient of variation of ≤5% were considered for genome size measurement, calculated as $2C_D = (C1/C2) \times 2C_S$, where $2C_D$ is the 2C value (pg) of each Cleomaceae accession, C1 is the mean G_0/G_1 peak channel of the Cleomaceae accession, C2 is the mean G_0/G_1 peak channel of *S. lycopersicum*, and $2C_S$ is 2.00 pg of *S. lycopersicum*. Based on previous results, the *S. lycopersicum* internal standard was replaced with *Raphanus sativus* 'Saxa' (2C = 1.13 pg, Praça-Fontes *et al.*, 2011) to measure the 2C value of *G. gynandra* (GG, Mossoró-RN) and *T. siliculifera* (TSI, Rio Pardo-MG).

Gas exchange and fluorescence analyses

Fully expanded leaves from the third node of non-flowering plants were used for gas exchange measurement with LICOR 6400XT (LI-COR Biosciences, Lincoln, USA). The measurements were performed at a flow rate of 400 µmol CO_2 s^{-1} and a light intensity of 1,000 µmol photons m^{-2} s^{-1}. All measurements were performed using a 2 cm^2 leaf chamber at 25°C. The leaf-to-air vapor pressure deficit was maintained at 1.2–2.0 kPa, and the amount of blue light was set to 10% PPFD to optimize the stomatal aperture. Chlorophyll *a* fluorescence parameters were quantified as described previously (Yin *et al.*, 2009). Corrections for CO_2 leakage into and water vapor from the leaf chamber of LI-6400 were applied to all gas exchange data, as previously described (Rodeghiero *et al.*, 2007). Dark respiration (R_d) was measured using the same gas exchange system described above, after incubation of at least 1 h in the dark. R_d was divided into two groups to estimate the mitochondrial respiration rate in light (R_L; Rodeghiero *et al.,* 2007).

Initial fluorescence (F_0) was measured by illuminating 1 h dark-adapted leaves with weakly modulated measuring beams (0.03 μmol photons $m^{-2} \cdot s^{-1}$). A saturating white light pulse (8,000 μmol photons $m^{-2} \cdot s^{-1}$) was applied for 0.8 s to obtain the maximum fluorescence (F_m), from which the variable-to-maximum chlorophyll *a* fluorescence ratio was calculated as $F_v/F_m = [(F_m - F_0)/F_m)]$. In light-adapted leaves, the steady-state fluorescence yield (F_s) was measured by applying a saturating white light pulse (8,000 μmol photons $m^{-2} \cdot s^{-1}$) to achieve light-adapted maximum fluorescence (F_m'). A far-red illumination (2 μmol photons $m^{-2} \cdot s^{-1}$) was applied after turning off the actinic light to measure the light-adapted initial fluorescence (F_0').

The capture efficiency of excitation energy by open photosystem II (PSII) reaction centers (F_v'/F_m') was estimated as described by Logan *et al.* (2007), and the actual PSII photochemical efficiency (Φ_{PSII}) was estimated as $(\Phi_{PSII}) = (F_m' - F_s')/F_m'$ (Genty *et al.*, 1989). As Φ_{PSII} represents the number of electrons transferred per photon absorbed by PSII, the electron transport rate (J_{flu}) was calculated as $J_{flu} = \Phi_{PSII} \cdot \alpha \cdot \beta \cdot \text{PPFD}$, where α is leaf absorptance and β is the partitioning of the absorbed quanta between PSII and PSI; the value of the product $\alpha \cdot \beta$ was adopted as described in the literature for *Arabidopsis* (0.451) (Flexas *et al.,* 2007).

CO_2 concentration at carboxylation sites (C_c) was calculated using the equation described by Harley *et al.* (1992), as follows:

$$C_c = (\Gamma^*(J_{flu} + 8(A_N + R_L))/(J_{flu} - 4(A_N + R_L)$$

The conservative value of Γ^* was adapted from an *Arabidopsis* study (Walker and Cousins, 2013). Mesophyll conductance (g_m) was estimated as the slope of the A_N versus $C_i - C_c$ relationship ($C_i < 300$ μmol·mol^{-1}). Given that the current methods for estimating g_m include several assumptions as well as technical limitations and sources of errors that need to be considered to obtain reliable values (Pons *et al.,* 2009), g_m was estimated using the Ethier and Livingston (2004) method, which involved the fitting of the A_N/C_i curve to the basic nonrectangular hyperbolic Farquhar–von Caemmerer–Berry (FvCB) model, based on the hypothesis that g_m reduces the curvature of the RuBisCo-limited portion of the A_N/C_i curve.

Metabolite analyses

Fully expanded leaves from the third node of non-flowering plants were harvested in the middle of the light period, soon after gas exchange and fluorescence analyses. At this point, the herbaceous plants were 2-month old and the other shrub species were 4-month old. The samples were snap-frozen and stored at -80°C until analysis. Metabolites were extracted by grinding the material in liquid nitrogen and adding the appropriate extraction buffers. Approximately 25 mg of the freeze-dried leaf matter was subjected to ethanolic extraction at 70°C for 30 min. After centrifugation (at 21,800 × *g* for 5 min), the chlorophyll content was determined as previously described (Porra *et al.,* 1989). The total chlorophyll content (*a*+*b*) and chlorophyll *a*/*b* ratio were calculated (Porra *et al.,* 1989).

The concentration of glucose, fructose, and sucrose was determined from the liquid phase, as described previously (Fernie *et al.,* 2001). Total free amino acid *(Cross et al.,* 2006), malate, and fumarate (Nunes-Nesi *et al.,* 2007) content was determined as previously described. From the ethanol insoluble fraction, the concentration of starch (Fernie *et al.,* 2001) and total soluble proteins (Bradford, 1976) was determined according to standard protocols.

Metabolite profiling

Fully expanded source leaves from each sample plant were collected in the middle of the light period, wrapped in aluminum foil, flash frozen in liquid nitrogen, and stored at -80°C until analysis. After lyophilization and grinding, approximately 15 mg dry weight was extracted for gas chromatography–mass spectrometry (GC–MS) in1.5 mL of extraction solution (water:methanol:chloroform, 1:2.5:1 *v/v*) as described by Fiehn (2007), using 5 µM ribitol (Sigma-Aldrich) as the internal standard. The extract (15 µL) was dried in a vacuum concentrator and derivatized in two steps using the MPS Dual Head autosampler (Gerstel), as described by Gu *et al.* (2012). After incubation for 2 h at room temperature, the samples were analyzed as described previously (Shim *et al.*, 2020) using the 5977B GC-MSD system (Agilent Technologies). The peaks were integrated using MassHunter Quantitative (v b08.00; Agilent Technologies). For relative quantification, metabolite peak areas were normalized to the corresponding dry weight and peak area of the internal standard (ribitol).

Carbon isotope composition analysis

Leaf material was oven-dried at 60°C for 48 h and used to determine $\delta^{13}C$ and the C/N ratio. The material was analyzed using the Isoprime 100 isotope ratio mass spectrometer coupled to an isotope cube elemental analyzer (Elementar, Hanau, Germany), as described by Gowik *et al.* (2011).

Micromorphological analyses

Anatomical analysis was conducted to evaluate the VD of fully developed leaves in the upper third. Diaphanization of the plant material was performed as described previously (Zsögön *et al.,* 2015). Subsequently, the samples were assembled on glass slides, and images of the adaxial epidermis were obtained using the Zeiss Axio Scope A1 photomicroscope coupled to a color image capture system (Axiovision® 105). Five photographs of each slide were obtained to cover the entire leaf length. In addition, five replicates (five leaves per individual) were collected and analyzed. Vein density was calculated as the length of the second-, third- and fourth-order veins in a given leaf area (mm mm^{-2}). Quantification was performed using Image Pro-Plus® (v 4.5; Media Cybernetics, Silver Spring, USA). In addition, the median region of the fully expanded leaves from the third node was collected and fixed in formaldehyde:acetic acid:ethanol 50% (FAA50) at a ratio of 5:5:9 (*v/v*) for 48 h. The material was dehydrated in an ethanolic series and embedded in historesin (Leica, Heidelberg, Germany). Transverse sections (thickness, 5 µm) were obtained using a self-advancing rotary microtome (RM 2155, Leica, Heidelberg, Germany) with glass razors. The sections were stained with 0.05% toluidine blue in acetate buffer (pH 4.7) (O'Brien *et al.*, 1964) for 1 min and assembled between slides and coverslips with synthetic resin. Photomicrographs were obtained using a photomicroscope (Zeiss, MC-80). The micrometric scales were photographed and enlarged under the same optical conditions. Leaf thickness was measured considering the entire length of the epidermis. For each slide analyzed, five measurements were obtained in different regions to cover the entire area. The distance between the vascular bundles was measured between the sides of the veins. Bundle sheath cell width (BSCW) was measured radially in five cells at random on each slide (cells from different vascular bundles were selected when possible). Leaf thickness, palisade parenchyma, spongy parenchyma, intervein distance, and BSCW were measured (Supplementary Figure S1) using Image Pro-Plus®.

Statistical analyses

Data were obtained from five plants per accession, each placed in an individual pot; thus, each pot represented a biological replicate. All plants were completely randomized. The effect of accession was determined using analysis of variance ($P < 0.05$). Means were analyzed using Tukey's test.

Phylogenetic generalized least squares (PGLS) regression was applied. We used a phylogenetic tree constructed using BI based on the ITS sequences of 21 novel Brazilian accessions and a matrix of 74 variables. It is noteworthy that the species *G. gynandra* (C_4 species) was removed from the analysis, as its genome size values (2-4 times greater compared to the other species in this study) and those obtained in the other analyzes could bias/influence the regression result. The script available at https://lukejharmon.github.io/ilhabela/instruction/2015/07/03/PGLS/ was used. PGLS uses the common statistical mechanism of generalized least squares and applies it to phylogenetic comparative data (Felsenstein, 1985). This method allows for the estimation of the impact of phylogeny on the covariance among residuals, thereby controlling for relatedness. In the case of Brownian motion, residues with variances and covariances follow the structure of the phylogenetic tree (Felsenstein, 1985). All statistical analyses were performed using Statistica and R.

RESULTS

Carbon isotope composition

The carbon isotope composition of plants has been widely used to identify the photosynthetic mechanisms, and it can, therefore, assist phylogenetic studies exploring the evolution of C_4 photosynthesis (Caemmerer *et al.*, 2014; Yang *et al.*, 2017). Hence, to identify the photosynthetic mechanism of the accessions studied here and to merge these data with molecular phylogeny, we performed carbon isotope composition ($\delta^{13}C$) analysis using fully expanded leaves of plants growing under the same conditions. In the Cleomaceae accessions tested, the $\delta^{13}C$ value ranged from -15‰ to -32‰, representing at least two types of photosynthetic mechanisms, namely C_3 and C_4 (Figure 1A). As expected, the Brazilian accession of *G. gynandra* exhibited the lowest $\delta^{13}C$ value (-15‰), and the other species showed values around -32‰. Meanwhile, *T. siliculifera* and *T. longicarpa* (Lavras, CE) exhibited $\delta^{13}C$ values close to -29‰, being significantly different from the other accessions (Figure 1A).

Phylogeny of Cleomaceae accessions

Our Bayesian phylogenetic reconstruction based on the ITS sequences of Cleomaceae species indicated that most of our analyzed accessions were positioned in the *Tarenaya* genus. The exceptions were the sequences of *C. paludosa* and *G. gynandra*, which were placed into the *Cleoserrata* and *Gynandropsis* clusters, respectively (Figure 1B). Moreover, *Tarenaya* is sister to two American genera, namely *Cleoserrata* and *Melidiscus*, whereas *Gynandropsis* is sister to the African cluster comprising *Coalisina* and *Iltisella* as well as to *Cleomella* (Figure 1B). These results are contrary to the findings reported by Bayat *et al.* (2018), who used five molecular markers (ITS, *mat*K, *ndh*F, *ycF*, and rps) and another set of species/sequences. Interestingly, *Gynandropsis*, *Areocleome*, and *Coalisina*, which harbor species with C_4 photosynthetic metabolism, formed different clusters (Figure 1B), indicating that this photosynthetic mechanism has evolved independently at least three times within the family, as was previously reported (Patchell *et al.*, 2014; Bayat *et al.*, 2018).

The analyzed genera, namely *Gynandropsis*, *Cleoserrata*, and *Tarenaya* are monophyletic. Within the *Tarenaya* group, the Rosea, Parviflorae, and Aculeatae series are monophyletic; the Spinosa series, which includes the type species of the genus (*T. spinosa*), is polyphyletic; and the Aculeatae series, which includes herbaceous species bearing small flowers, is placed more externally to the other series of the genus (Figure 1B). Within the Spinosa series, *T. hassleriana* accessions form different sub-clusters, indicating high genetic diversity among the species (Figure 1B). Overall, the 20 accessions of *Tarenaya* studied cover a great genetic diversity of genus, mainly because it includes species from all series.

Genome size in Cleomaceae

In flow cytometry histograms with G_0/G_1 peaks of nuclear suspensions, the coefficient of variation for all Cleomaceae species/accession and the internal standards (*S. lycopersicum* or *R. sativus*) was <3.70%. Based on genome size, the 22 accessions were separated into seven groups: (i) *T. hassleriana* accessions (THC, THCS, THDM, THJ, THP, THS, and THV) with 2C = 0.55 pg ± 0.003; (ii) *T. diffusa* (TD) with 2C = 0.59 pg ± 0.004; (iii) *T. aculeata* and *T. microcarpa* (TA and TM) with 2C = 0.66 pg ± 0.004; (iv) *T. siliculifera* (TSI) with 2C = 0.77 pg ± 0.003; (v) *C. paludosa* (CP) with 2C = 1.08 pg ± 0.004; (vi) nine *Tarenaya* species (TS,

Table 1. Description of the novel Brazilian species used for molecular phylogeny analyzes, and their respective collection sites (city-state abbreviation) and biomes.

Species	ID[#]	Habit	City – State*	Local - Environment	Biome
Cleoserrata paludosa	CP	sub-shurb	Belém – PA	marshy region	Amazon Forest
Gynandropsis gynandra	GG	herb	Mossoró – RN	crop invasive	Cerrado - Savana
Tarenaya aculeata	TA	herb	Feira de Santana – BA	open field, abandoned	Atlantic Forest
T. diffusa	TD	herb	Feira de Santana – BA	open field, abandoned	Atlantic Forest
T. hassleriana	THV	shurb	Viçosa – MG	marshy region	Atlantic Forest
T. hassleriana	THC	Shurb	Canaã – MG	open field, abandoned	Atlantic Forest
T. hassleriana	THD	Shurb	Domingos Martins – ES	roadside	Atlantic Forest
T. hassleriana	THS	Shurb	São Miguel– MG	open field, abandoned	Atlantic Forest
T. hassleriana	THSC	Shurb	Canoinhas - SC	open field, abandoned	Atlantic Forest
T. hassleriana	THP	Shurb	Piau - MG	roadside	Atlantic Forest
T. hassleriana	THJ	Shurb	Joinville – SC	roadside	Atlantic Forest
T. longicarpa	TL	Shurb	Picos – PI	open field, abandoned	Caatinga - Savana
T. microcarpa	TM	Herb	Belém – PA	marshy region	Amazon Forest
T. parviflora	TP	Shurb	Pombal – PB	marshy region	Caatinga - Savana
T. rosea	TR	Shurb	Colatina – ES	open field, abandoned	Atlantic Forest
T. longicarpa	TAM	Shurb	Manaus – AM	open field	Amazon Forest
T. longicarpa	TARC	Shurb	Arcoverde - PE	open field, abandoned	Caatinga – Savana
T. longicarpa	TIB	Shurb	Ibimirim - PE	open field, abandoned	Caatinga – Savana
T. longicarpa	TAF	Shurb	Afrânio - PE	open field, abandoned	Caatinga – Savana
T. longicarpa	TC	Shurb	Lavras do Ceará - CE	open field, abandoned	Caatinga – Savana
T. spinosa	TS	Shurb	Teresina - PI	roadside	Caatinga - Savana
T. siliculifera	TSI	shurb	Rio Pardo - MG	altitude fields	Atlantic Forest

[#] Acronyms used for species identification. *States: PA – Pará; RN – Rio Grande do Norte; BA – Bahia; MG – Minas Gerais; ES – Espírito Santo; SC – Santa Catarina; PI – Piauí; PB – Paraíba; AM – Amazonas; CE – Ceará; PE – Pernambuco.

TP, TL, TR, TAF, TAM, TARC, TC, and TIB) with 2C = 1.30 pg ± 0.014; and (vii) *G. gynandra* (GG) with 2C = 2.20 pg ± 0.002 (Figure 1B; Supplementary Figure S2). Thus, we observed interspecific variation in nuclear genome size among the Brazilian Cleomaceae species. Based on the mean 2C values and phylogeny (Figure 1B), we speculate that increase and decrease in genome size contributed to the diversification of Cleomaceae species, specifically the C_4 species *G. gynandra* (2C = 2.20 pg, four-fold higher than that of *T. hassleriana*). Considering the correlation between the 2C value and the 2n chromosome number of *Tarenaya* species and *G. gynandra*, we speculate that euploidy and disploidy (ascendant and/or descendant) resulted in interspecific variations in the 2C value and C_4 photosynthetic mechanism of *G. gynandra*. Alternatively, genome size differences may be attributable to whole genome duplications (Mabry *et al.,* 2020) and/or changes in repetitive DNA content (Beric *et al.,* 2020; 2021).

Furthermore, PGLS regression was applied to determine the correlations of genome size with the other study variables, independently, in the phylogenetic topology of the Brazilian

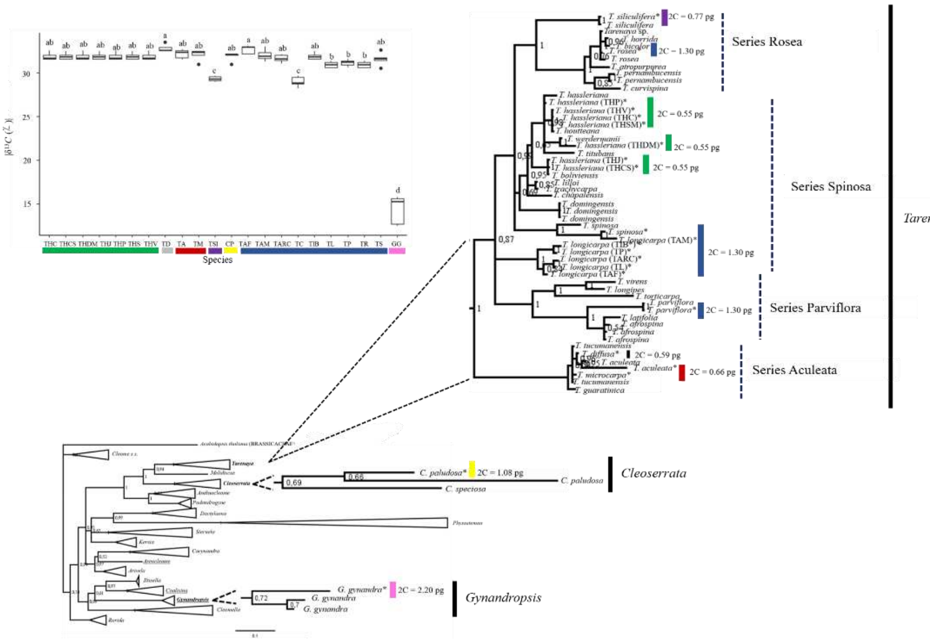


Figure 1. Carbon isotopic discrimination and molecular phylogeny of Cleomaceae species. (A) Carbon isotope composition (δ^{13}C; ‰). Letters above individual box-scatter indicate significant groupings according to Tukey's Test (n=5). The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. Morphospecies studied. THC: *T. hassleriana* (Canaã-MG); THCS: *T. hassleriana* (Canoinhas-SC); THDM: *T. hassleriana* (Domingos Martins-ES); THJ: *T. hassleriana* (Joinville-SC); THP: *T. hassleriana* (Piau-MG); THS: *T. hassleriana* (São Miguel-MG); THV: *T. hassleriana* (Viçosa-MG); TD: *T. diffusa* (Feira de Santana-BA); TA: *T. aculeata* (Feira de Santana-BA); TM: *T. microcarpa* (Belém-PA); TSI: *T. siliculifera* (Rio Pardo-MG); CP: *C. paludosa* (Belém-PA); TAF: *T. longicarpa* (Afrânio-PE); TAM: *T. longicarpa* (Manaus-AM); TARC: *T. longicarpa* (Arcoverde-PE); TC: *T. longicarpa* (Lavras-CE); TIB: *T. longicarpa* (Ibimirim-PE); TL: *T. longicarpa* (Picos-PI); TP: *T. parviflora* (Pombal-PB); TR: *T. rosea* (Colatina-ES); TS: *T. spinosa* (Teresina-PI); GG: *G. gynandra* (Mossoró-RN). The acronyms are followed by the species name and city/state of sampling state, between parenthesis. The coloured bars are related to the groups observed in Figure 1B. **(B)** Bayesian Inference consensus tree inferred from nuclear ribosomal ITS sequences retrieved from Cleomaceae and Brassicaceae members, demonstrating the phylogenetic relationship between Cleomaceae genera. Underlined sequences represent genera that have species with C_4 photosynthetic metabolism (*Gynandropsis, Coalisina* and *Areocleome*). In bold are the genera analyzed in this study (*Tarenaya*, *Cleoserrata* and *Gynandropsis*). The sequences produced in this study are marked with an asterisk. In addition, the 2C nuclear value for each species (group) was added following its identification. The colors represent different genome sizes. Numbers at nodes reflect PP. Bar: 0.2 nucleotide substitutions per site.

Cleomaceae accessions (Figures 2-3). It is noteworthy that a phylogeny, based on the studied species (except *G. gynandra*) was made exclusively for PGLS. Thus, with a smaller sample of species and a single molecular marker (ITS), it is possible to see that the *Tarenaya* series were no longer monophyletic (Supplementary Figure S3).

In PGLS regression 32 of the 74 variables were significantly correlated to genome size ($P < 0.05$) (Figures 2-3). Based on the significant regressions, the models with, photosynthetic rate (A_N), A_{gross}, palisade parenchyma, venation density, Respiration rate (R_D), malate, aspartate, clorophyll $a + b$, glucose, α-alanina, putrescine, methionine, phenylalanine, tyrosine, glycerol, glutamine, serine, ornithine, Gaba, glycine, gluconate, isocitrate and leucine were positively correlated with genome size. However, some important variables, such as the $\delta^{13}C$ value, intrinsic water use efficiency (WUEi), malonate, myinositol, proline, starch content, proteins content, seeds weight and SLA were negatively correlated with genome size (Figures 2-3).

Physiological parameters

To completely characterize and understand the physiological basis for the differences among the studied Cleomaceae accessions, we performed detailed physiological analyses of diffusional, photochemical, and biochemical constraints on photosynthesis (Figure 4A-D; Supplementary Figure S4). Under the tested environmental conditions (as described in the Material and Methods section), Cleomaceae accessions showed natural variation, even within the same species, in terms of physiological parameters related to photosynthesis. In general, the studied Cleomaceae accessions exhibited a high A_N, with values greater than 20 µmol CO_2 $m^{-2}\cdot s^{-1}$ (Figure 4A). However, the herbaceous accessions of *T. aculeata*, *T. diffusa*, *and T. microcarpa* exhibited the A_N of ~10 µmol CO_2 $m^{-2}\cdot s^{-1}$. while the shrubby accession of *T. siliculifera* showed the A_N of <10 µmol CO_2 $m^{-2}\cdot s^{-1}$. *T. siliculifera* exhibited the lowest stomatal conductance (g_s), followed by *G. gynandra*, herbaceous species (*T. aculeata*, *T. diffusa*, and *T. microcarpa*), and *Tarenaya* species (Spinosa I and II, Parviflorae, and Cleorosea) (Supplementary Figure S4). For the internal CO_2 concentration (C_i), which is related to g_s, the same pattern as that for g_s was observed (Supplementary Figure S4). Regarding transpiration (E) and R_d, the studied Cleomaceae accessions exhibited substantial differences, and we, thus, did not observe any pattern between groups (Figure 4; Supplementary Figure S4). Similarly, WUE_i was remarkably higher in *G. gynandra*, albeit without a clear trend in the rest of the species (Figure 4). We further estimated several photosynthesis-related parameters, such as the maximum rate of RuBisCO carboxylase activity (V_{cmax}) and electron transport rate (J_{flu}) (Supplementary Figure S4). Consistent with the other parameters, we observed a large variability in these two traits. The highest V_{cmax} was observed for *T. longicarpa* (Arcoverde, PE), and the highest J_{flu} was observed for *T. hassleriana* (Canaã, MG), *T. hassleriana* (Viçosa, MG), and *T. longicarpa* (Afrânio, PE) (Supplementary Figure S4).

Biochemical analyses

To unveil the natural variations among the different Cleomaceae accessions studied, we performed detailed metabolite analyses using leaf samples. First, we examined the levels of the major carbon- and nitrogen-containing metabolites in leaves harvested at the middle of the light period. Overall, despite clear variations in physiological traits (Figure 4E-F; Supplementary Figure S5), the steady-state levels of soluble sugars (glucose, fructose, and sucrose), starch, amino acids, proteins, and chlorophylls did not differ among the plants of different species (Figure 4E-F, Supplementary Figure S5). The only exception to this trend was *T. siliculifera*, which exhibited higher steady-state levels of sugars and starch than the other species (Figure 4E, Supplementary Figure S5). Meanwhile, *G. gynandra* exhibited higher levels of amino acids (Figure 4E) and proteins than *T. diffusa* (Figure 4F). There were no significant differences in the chlorophyll *a*/*b* ratio among the studied species (Supplementary Figure S5).

Furthermore, we performed metabolite profiling of leaf samples from all studied accessions using GC–MS. This analysis allowed for the identification and quantification of 33 metabolites (Supplementary Table S2, Figure S6). For 19 metabolites analyzed, we did not observe significant differences among the studied accessions (Supplementary Figure S6; Table S2). Among these metabolites, there were no differences in the levels of the amino acids aspartate, glycine, and serine or the organic acids citrate and isocitrate. In contrast, *T. siliculifera* exhibited significantly higher levels of myo-inositol, succinate, malonate, malate, threonate, glycerate, glycerol, glucose, fructose, and sucrose. In addition, *T. hassleriana* exhibited higher levels of methionine, glutamate, and alanine.

Leaf anatomy

In leaves, the formation of venation, plasmodesmata, and other barriers to gas diffusion is important for photosynthetic efficiency. Thus, the variability in leaf anatomical traits, which are related to the carbon concentrating mechanism, was determined in the Cleomaceae accessions studied. Leaf anatomical parameters were evaluated by diaphanization and transverse sectioning to determine VD and BSCW. *T. hassleriana* (THC, THCS, THP, and THS), *T. longicarpa* (TP, TC, and TIB), *T. rosea*, and *T. spinosa* showed higher VD than the other analyzed accessions (Figure 4G). Meanwhile, *T. hassleriana* (THDM and THV), *T. longicarpa* (TAF, TAM, and TARC), and *T. parviflora* showed intermediate VD, whereas *G. gynandra*, *T. hassleriana* (THJ), *T. diffusa*, *T. aculeata*, *T. microcarpa*, *T. siliculifera*, and *C. paludosa* exhibited lower VD (Figure 4G).

BSCW was higher in *G. gynandra*, *T. hassleriana* (THC and THCS), *T. siliculifera*, and *T. longicarpa* (TAM) than in the other accessions studied (Supplementary Figure S7).

Meanwhile, *T. diffusa, T. aculeata*, *T. microcarpa*, *T. parviflora*, and *T. longicarpa* (TARC and TAF) exhibited lower BSCW. The remaining 11 accessions exhibited intermediate BSCW (Supplementary Figure S7).

The interveinal distance (ID) was higher in *T. microcarpa* but lower in *G. gynandra* and *T. siliculifera*. The other accessions showed intermediate ID (Supplementary Figure S7). *T. hassleriana* (THP) showed the highest leaf thickness, whereas *T. diffusa* and *T. microcarpa* showed the lowest leaf thickness. The other accessions exhibited intermediate values and did not differ in terms of thickness (Supplementary Figure S7). Regarding the thickness of the palisade (Figure 4H) and spongy parenchyma (Supplementary Figure S7), *T. hassleriana* (THP) exhibited higher values and *G. gynandra* showed lower values than the remaining accessions, which showed intermediate values (Supplementary Figure S7).

Surprisingly, we could not separate the accessions into groups based on these morphoanatomical characteristics, which would be expected considering that the C_4 species *G. gynandra* could be clearly distinguished from the other C_3 species in terms of most of the analyzed parameters (*e.g.*, VD and BSCW) (Supplementary Figure S8, S9). In this context, our results suggest that the analyzed accessions diversified at the species level. Since *T. hassleriana* and *T. longicarpa* differed in terms of the analyzed parameters, anatomy would provide additional evidence of this diversification. Although these traits indicate the possible diversification in photosynthetic mechanisms, physiological and biochemical data provide additional evidence.

DISCUSSION

Interspecific variations in the 2C value complement the molecular phylogeny of Cleomaceae

Cleomaceae is a highly diverse family, both morphologically (*e.g.*, monoecious, dioecious, polygamous species; *e.g.*, Omondi *et al.,* 2017; Riaz *et al.,* 2019; Zohoungbogbo *et al.,* 2018) and physiologically (species with different photosynthetic metabolisms: C_3, C_3-C_4, and C_4). This diversity was represented by species studied here, which were morphologically (*Tarenaya*, *Cleoserrata*, and *Gynandropsis*; Stevens 2001; Flora of Brazil 2021) and physiologically (*e.g.*, *Tarenaya* and *Gynandropsis*; Marshall *et al.,* 2007; Bayat *et al.,* 2018) distinct. The three genera analyzed here, namely *Tarenaya*, *Cleoserrata*, and *Gynandropsis*, were monophyletic (Figure 1B). Of note, however, *Gynandropsis* is a monotypic genus (*G. gynandra*). *Tarenaya*, one of the most diverse genera in terms of the number of species within Cleomaceae, also includes many morphologically distinct species, which allows for the separation of the genus into series

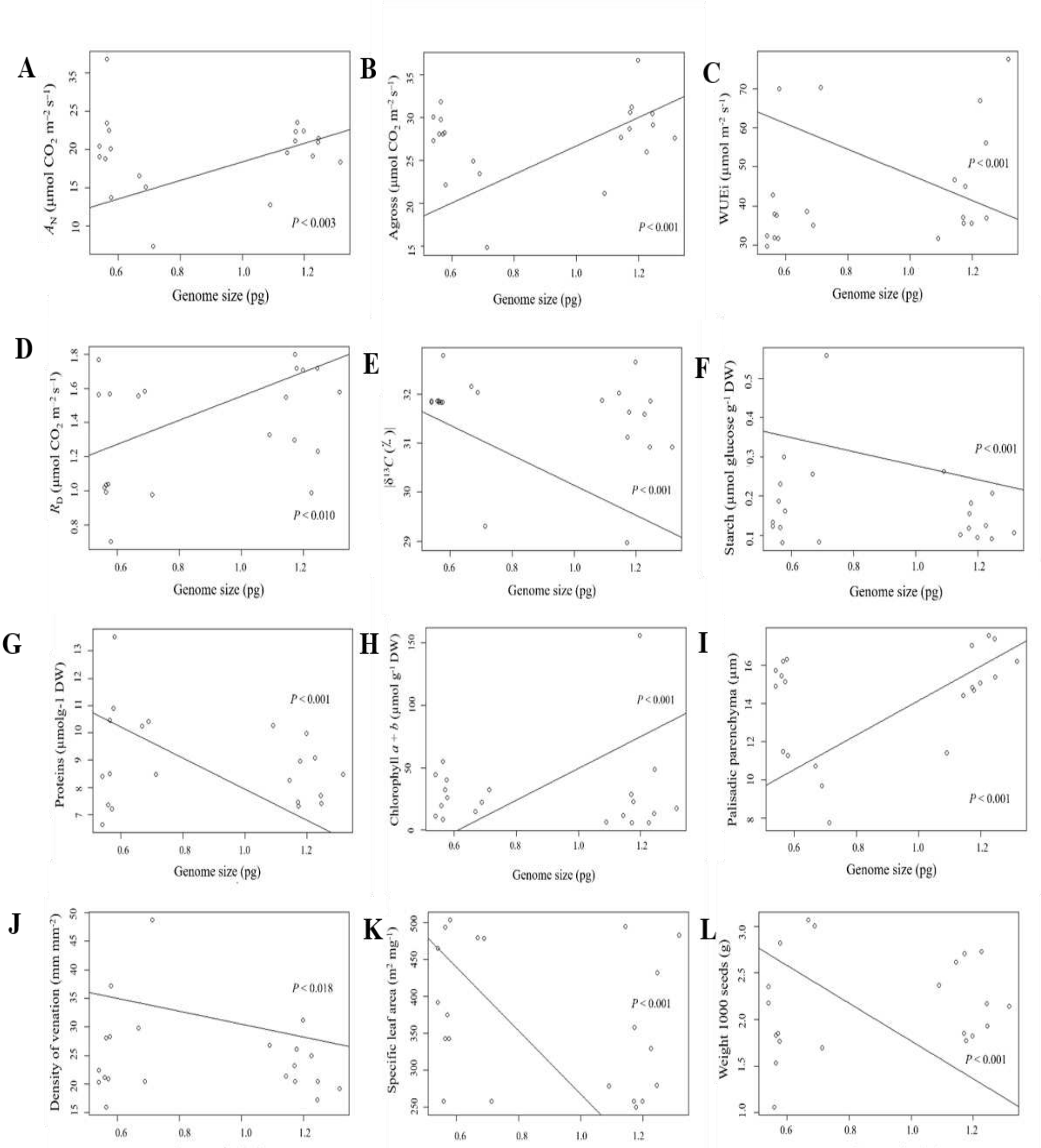


Figure 2. Evolutionary regression obtained from PGLS analysis (Brownian motion) conducted with 21 species. Only the variables that presented significant *P*-value ($P < 0.05$) are shown. Physiological variables: **(A)** Ambient CO$_2$ assimilation rates (A_N). **(B)** Agross. **(C)** intrinsic water use efficiency (WUEi). **(D)** Respiration (R_D). Metabolic variables: **(E)** isotopic carbon discrimination (δ^{13}C. ‰). Metabolic variable: **(F)** Starch content. **(G)** Proteins content. **(H)** chlorophyll *a* + *b* content. Anatomical variable: **(I)** Palisade parenchyma. **(J)** Density of venation. Growth parameter: **(K)** Specific leaf area. **(L)** Weight 1000 seeds.

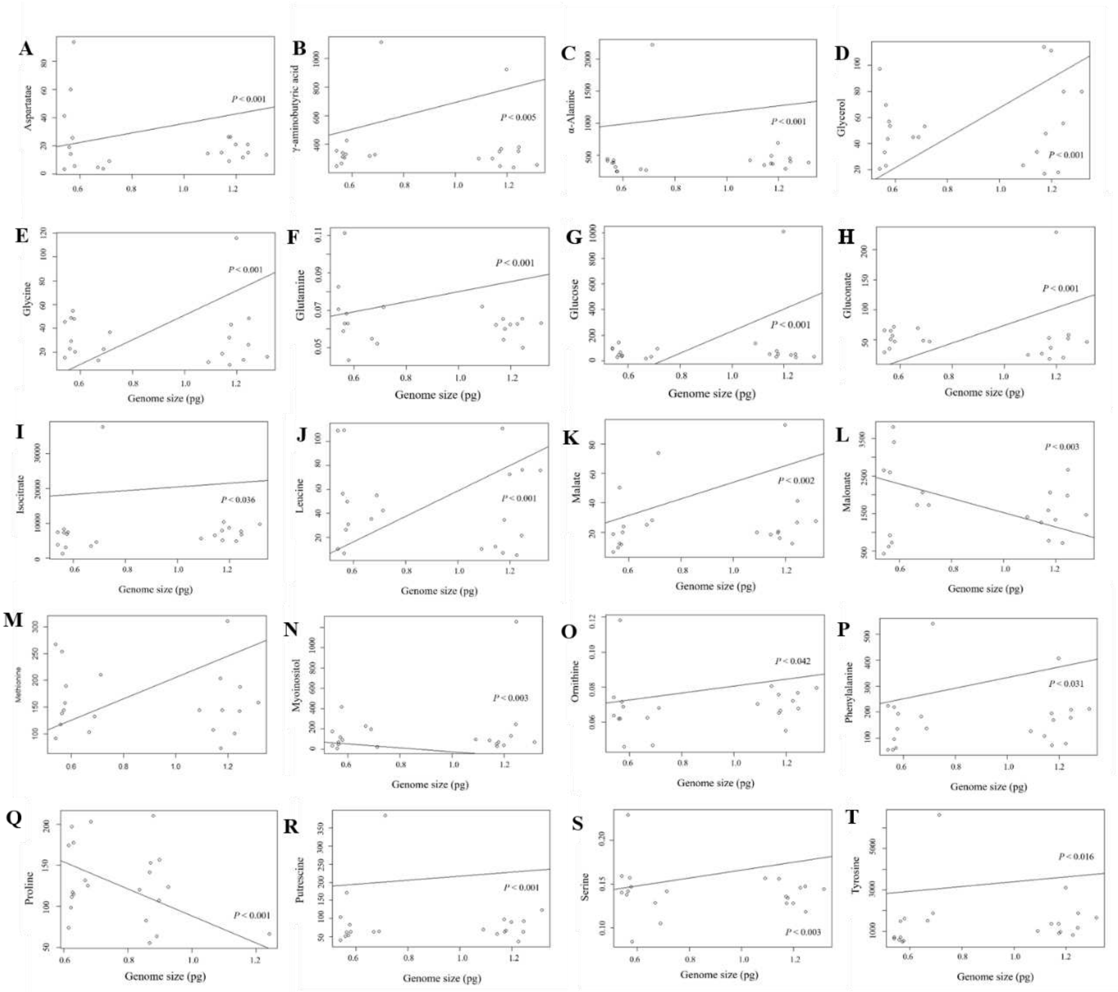


Figure 3. Evolutionary regression obtained from PGLS analysis (Brownian motion) conducted with 21 species. Only the variables that presented significant *P*-value (*P* <0.05) are shown. Relative metabolite content from GC-MS**.** **(A)** Aspartate. **(B)** γ-aminobutyric acid. **(C)** α-alanine. **(D)** Glycerol. **(E)** Glycine. **(F)** Glutamine. **(G)** Glucose. **(H)** Gluconate. **(I)** Isocitrate. **(J)** Leucine. **(K)** Malate. **(L)** Malonate. **(M)** Methionine. **(N)** Myoinositol. **(O)** Ornithine. **(P)** **(Q)** Proline. **(R)** Putrescine. **(S)** Serine. **(T)** Tyrosine.

(small groups of species) (*e.g*., Iltis, 1957, 1959; Hall *et al.,* 2002; Sanchez-Acebo, 2005; Inda *et al.,* 2008; Feodorova *et al.,* 2010; Iltis and Cochrane, 2014). These series have previously been shown to be monophyletic (see Soares-Neto *et al.*, 2020), and our results are largely congruent with this report, with the exception of the Spinosa series, which we found to be polyphyletic. Of note, our study is limited by the number of samples of this rather diverse series (see Soares-Neto *et al.*, 2020). Thus, studies with increased taxon sampling or markers are warranted to resolve the monophyly of the Spinosa series.

The morphological and physiological diversity can be linked to variations in genome size, among other factors (Paterson *et al.,* 2010; Roddy *et al.,* 2019). Thus, an interspecific variation at the diploid level, for example, as found for the species *T. hassleriana, T. diffusa, T. aculeata* and *T. microcarpa*, may suggest that the DNA content may be a parameter that can be used to differentiate the species (*e.g.*, *Miscanthus* 2C = 3.9 to 6.9 pg – Sheng *et al.,* 2016). However, the observed species and genome size diversity was not consistent with the molecular phylogenetic results (Figure 1B; Supplementary Figure S2; Patchell *et al.,* 2014; Bayat *et al.,* 2018). Different, for example, from the one observed for the genus *Lupinus* (Fabaceae), in which the genome size (0.97 to 2.44 pg) data supports the generally accepted taxonomic classification of the Old World lupins (Naganowska *et al.,* 2003). In this sense, we observed that different series have species with the same genome size (Parviflora, Rosea and Spinosa). Furthermore, within the same series it is also possible to observe representatives with different genome sizes (*e.g. T. hasseriana* and *T. spinosa* - Spinosa series, and *T. rosea* and *T. siliculifera* - Rosea series). It is noteworthy that the species of Cleomaceae studied have an average genome size (0.5 – 2.2 pg), larger than that recorded for *Arabidopsis*, for example, but corresponding to most species that have their genome quantified (Bai *et al.*, 2012), such as such as rice, tomatoes and fruit – such as cherry, mango, papaya, orange and peach (Arumuganathan and Earle 1991).

Polyploidy and/or the expansion of repetitive DNA content have played a major role in the evolution of many species by increasing the genome size and gene copy number (Schranz and Mitchell-Olds, 2006; Wang *et al.,* 2009; Qiao *et al.,* 2019; Tao *et al.*, 2019; Roddy *et al.,* 2020; Marbry *et al.,* 2020; Beric *et al.,* 2020)· For instance, we showed that the genome size of *G. gynandra* is two to four times greater than that of some *Tarenaya* species (Figure 1; Supplementary Figure S2). As some *Tarenaya* species have 2n = 18 chromosomes (*e.g.*, *T. hassleriana* is considered diploid; Inda *et al.,* 2008) and *G. gynandra* has 2n = 34 chromosomes (considered tetraploid, Omondi *et al.,* 2017), polyploidy seems to have increased in genome size in the latter. However, for narrowly defined phylogenetic groups (*e.g.*, genera), genome size and ploidy are positively correlated (Leitch and Bennett 2004; Dodsworth *et al.,* 2016). As such, the Spinosa series of *Tarenaya* (Inda *et al.,* 2008; Patchell, Roalson, and Hall 2014) includes tetraploid species (*e.g.*, *T. spinosa and T. longicarpa*), while most other species of the genus (*e.g.*, *T. aculeata*, *T. diffusa*, and *T. microcarpa*), and even some species of the Spinosa series (*e.g.*, *T. hassleriana*), are diploid (Inda *et al.,* 2008).

Based on the variations in genome size among Cleomaceae genera and even among species of the same genus (*e.g.*, *Tarenaya*; Figure 1B), even our small sample of the family reflects the relatively large interspecific variability. The intraspecific variations in the ITS sequence as well as the interspecific variations in this sequence and the 2C value may be related to different selective environmental pressures. Polyploidy, disploidy (ascendant and/or descendant), and genome size are associated with life-history traits, including vegetative form,

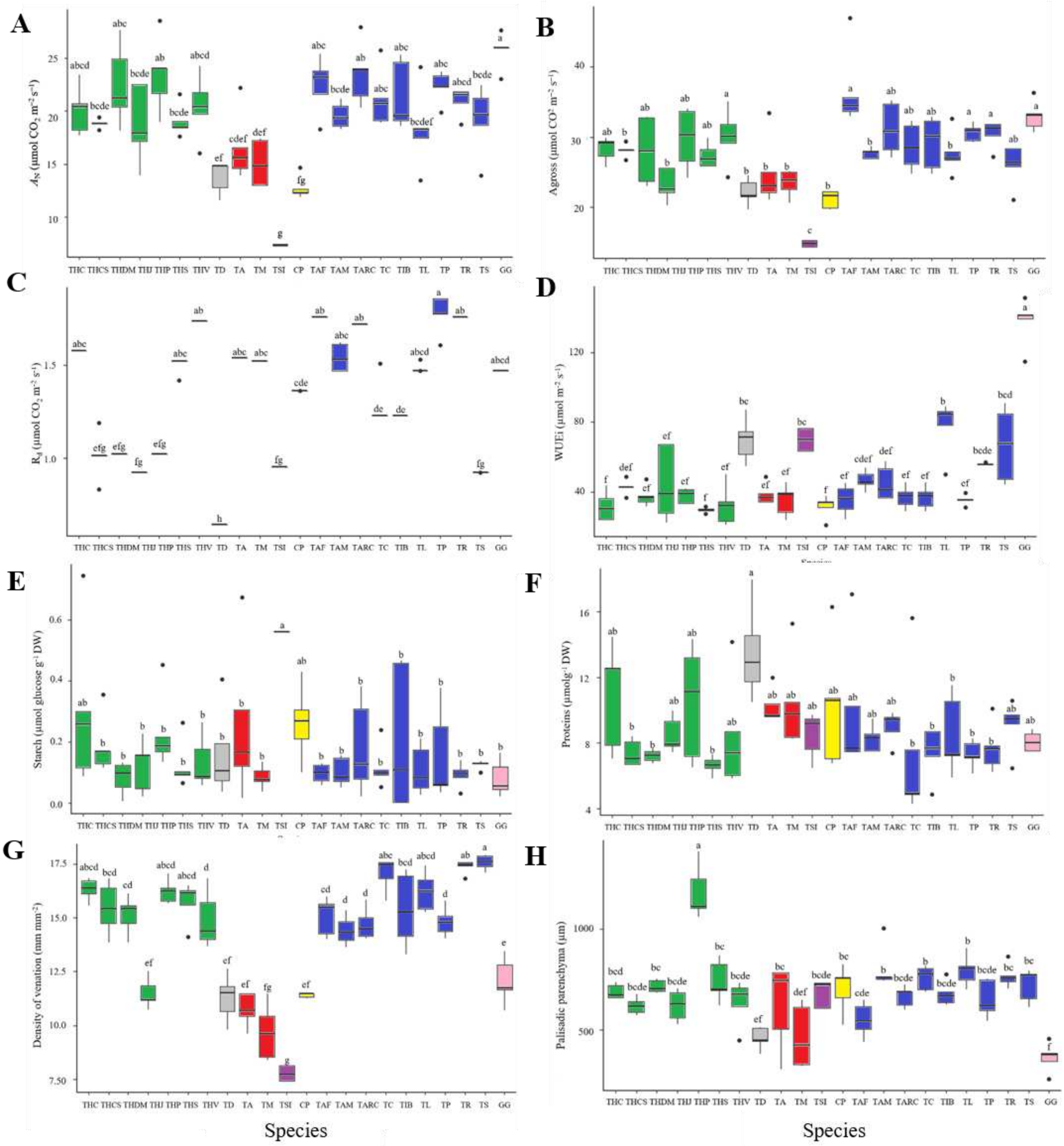


Figure 4. Gas exchange and chlorophyll *a* fluorescence parameters in Cleomaceae species. **(A)** Ambient CO_2 assimilation rates (A_N) (400 ppm atmospheric [CO_2]). **(B)** stomatal conductance (g_s). **(C)** dark respiration (R_d). **(D)** intrinsic water use efficiency (WUEi). Metabolic variable: **(E)** Starch content. **(F)** Proteins contents. Anatomical variable: **(G)** Density of venation. **(H)** Palisade parenchyma. Letters above individual box-scatter indicate significant groupings according to Tukey's Test (n=5). The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. The coloured bars are related to the groups observed in Figure 1 as well as the species acronyms.

flowering time, and adaptation to particular ecological niches, generating a basis for evolutionary novelty (Gregory 2002; Chris Pires *et al.,* 2006; Defoort *et al.*, 2019). However, only a few Cleomaceae species have been characterized in terms of their chromosome number (Subramania and Susheela 1988; Ammal, 1933; Inda *et al.,* 2008) and/or genome size (Omondi *et al.*, 2017), and further efforts are required to understand the karyotype evolution in these taxa.

In Brassicales, several whole genome duplication events (Mabry *et al.,* 2020; Beric *et al.,* 2021) have occurred, although no significant changes have been identified in the transposon landscape following these events (Beric *et al.,* 2020). Nuclear genome size can be used to understand the evolution of photosynthetic mechanisms (Bianconi *et al.,* 2018; Qiao *et al.,* 2019), as the increase and decrease in nuclear DNA content are the outcomes of karyotype alterations and are, thus, related to phenotypic changes (Sattler *et al.*, 2016). In plants, changes in genome size result from different karyotype events, particularly euploidy, aneuploidy, structural chromosome rearrangements, and mobile elements (transposons and retrotransposons). These changes can result in gene copy number modifications (Schranz and Mitchell-olds, 2006; Beric *et al.,* 2020), and new physiological, morphological, and reproductive phenotypes can be established, playing relevant roles in the diversification and speciation (Baltisberger and Hörandl, 2016) as well as the geographical distribution of the species (Tuler *et al.,* 2019).

The genome duplication contributed in Cleomaceae, for example, to several aspects of the evolution of C_4 photosynthesis in *G. gynandra* (Huang *et al.*, 2021). In this sense, the species has retained the duplicates of alanine aminotransferase and glutamine oxoglutarate aminotransferase; besides almost all known vein-development-related paralogous genes derived from the genome duplication event; the genome duplication also facilitated the evolution of C_4 enzyme genes and their recruitment into the C_4 pathway; in addition, several genes encoding photosystem I proteins were derived from the genome duplication (for more details see Huang *et al.,* 2021). Furthemore, there are some examples of the duplication of gene copy number and its outcomes in the evolution of C_4 metabolism, such as the phosphoenolpyruvate carboxylase kinase gene families and the carbonic anhydrase genes (Wang *et al.,* 2009). However, the number of enzymes required for C_4 metabolism is limited. In addition, the duplicated genes and other genomic regions often produce different effects (Lynch and Conery, 2003) and do not always result in functional innovation (Nielsen *et al.*, 2005). In this context, some studies have proposed the involvement of parallel evolution in changes of gene expression and amino acid sequences (Christin *et al.,* 2012; Williams *et al.,* 2012). Therefore, the cis-elements that direct cell specificity in C_4 leaves are present in C_3 orthologous genes recruited to the C_4 pathway, probably facilitating parallel evolution (Williams *et al.,* 2012).

The 2C value and evolution of C_4 photosynthetic mechanism in Cleomaceae

The increase and/or decrease in DNA content can promote phenotypic changes in the biochemical, anatomical, and physiological aspects of photosynthesis (Warner and Edwards, 1993; Vyas *et al.,* 2007; Oa *et al.,* 2011; Roddy *et al.,* 2020). The importance of species relationships in trait-based studies is well appreciated in the field of comparative biology (Guy *et al.,* 2019). Accordingly, *G. gynandra*, the species with C_4 photosynthesis, possesses a larger genome (Figure 1; Supplementary Figure S2). Nonetheless, a critical component in the study of adaptations is the identification of evolutionary correlations between phenotypic characteristics in phylogenetically close species (Adams, 2014). In this context, phylogenetic regression analysis provides a flexible analytical tool for assessing the degree of evolutionary association between variables while accounting for phylogeny (Adams, 2014).

In addition, the species with intermediate genome size (2C = 1.30 pg) (Figure 1B) had shorter ID and BSCW values, similar to the C_4 species *G. gynandra* (Figure 4; Supplementary Figure S7). The lowest ID observed in *G. gynandra* can be attributed to the decline in the mesophyll (M) cell number, as verified, for instance, by Marshall *et al.* (2007) and Reeves *et al.* (2018), and size, which are considered essential for optimal C_4 function, since fewer M cells reduce the mean diffusion distance for C_4 metabolites (Hattersley, 1984; Khoshravesh *et al.,* 2020). In addition, the enlargement of BSCs (equivalent to that in C_4 plants) is considered essential for C_4 function; as such, larger BSCs allow for a greater organelle volume and create larger vacuoles, serving as a barrier to facilitate CO_2 trapping in the sheath tissues (Von Caemmerer and Furbank, 2003; Khoshravesh *et al.*, 2020). Consistent with previous reports, our results indicated that increase in the genome size affected leaf morphological and anatomical traits, in addition to physiological function (Roddy *et al.,* 2020). Although the mechanism underlying the effect of genome size on cellular morphology remains unknown (Roddy *et al.,* 2020), its effects on the vascular transport network, vessel size, and density have already been demonstrated (Maherali *et al.*, 2009; Hao *et al.*, 2013; De Baerdemaeker *et al.*, 2018; Simonin and Roddy, 2018). In addition, genome size is a better predictor of guard cell size and stomatal density (Simonin and Roddy, 2018). Therefore, genome size can act as a first-order restriction on carbon gain, which directly affects the upper limit of allocation for growth, reproduction, and defense (Roddy *et al.,* 2020).

Based on this premise, genome size may affect the tolerance of environmental changes related to latitude, altitude, temperature, precipitation, salinity, and desiccation (Grime and

Mowforth, 1982; Otto and Whitton, 2000; Díez *et al.*, 2013; Wang *et al.*, 2018; Silva Artur *et al.*, 2019; Zhang *et al.*, 2019; Meyerson *et al.*, 2020). Therefore, in habitats that can support high productivity and primary metabolic rates (*e.g.*, Atlantic Forest and Amazon rainforest), species with smaller genomes (*e.g.*, *T. hassleriana*, *C. paludosa*, and *T. microcarpa;* Table 1) predominate, as they can maintain higher rates of metabolism and can more rapidly adjust their physiology to match the environmental conditions (Simonim and Roddy 2018; Roddy *et al.,* 2020). Since, the genome size restricts the minimum cell size and the maximum cell compaction densities, thus it can affect the maximum rate of photosynthetic metabolism in plants (Roddy *et al.,* 2020), in the absence of limiting factors. In this sense, species with smaller genomes there was greater variation in the sizes and packing densities of veins and stomata among which, consequently, allowed higher maximum photosynthetic rates (Simonim and Roddy 2018; Roddy *et al.,* 2020). Meanwhile, arid habitats (*e.g.*, Cerrado, Brazilian savanna, and Caatinga, Brazilian semi-arid region; Table 1) are characterized by low productivity and can support species with large genomes (Roddy *et al.,* 2020), such as *G. gynandra*, *T. spinosa*, and *T. longicarpa*, since a larger genome may be associated with greater greater genetic diversity (heterozygosity).

Although we speculate that genome size can predict the metabolic rate, its effects may be probably more nuanced, since numerous studies that have correlated the evolution of genome size with ecological factors have often produced conflicting results, and this correlation may change according to the group analyzed. Notably, studies addressing the diversity of genome size among closely related species and the association of genome size with phenotypes and ecological factors are scarce. In this light, the present study is the first to address variations in genome size within Cleomaceae, and further research is warranted to draw conclusions.

Are *Tarenaya* species on an evolutionary trajectory toward the C_3-C_4 photosynthetic mechanism?

C_4 photosynthesis arose from C_3 photosynthesis through a series of events/phases (Sage 2004; Gowik and Westhoff 2011). Combined physiological, molecular, and morphological changes played crucial roles in the evolution of the C_4 photosynthetic mechanism. The Cleomaceae accessions studied here were collected from warm and humid regions (Table 1). For instance, *C. paludosa*, *T. microcarpa*, and most *T. hassleriana* accessions were collected from riverbanks in abandoned fields, except *T. hassleriana* (Joinville, SC), which was on an access road. The other species studied were collected from abandoned fields (*T. longicarpa* or *G. gynandra*). Similar to the species collected in the present study, some species with intermediate metabolism (*e.g.*, *Flaveria linearis*; Holaday *et al.,* 1984) grow in comparable environments and at disturbed sites, such as abandoned fields and roads. Thus, taking the steps essential for the development of the C_3-C_4 and C_4 photosynthetic mechanisms into account (see Sage 2004; Gowik and Westhoff 2011) and given that the studied Cleomaceae species occur in favorable environments (Table 1) for the development of the C_3-C_4 photosynthetic mechanism, our results suggest that *Tarenaya* species as well as *Cleoserrata* may be pre-conditioned to evolve the characteristics associated with the C_4 photosynthetic mechanism. Similar trends have been observed in a previous study on *Cleome foliosa* and *Cleome africana* (Marshall *et al.,* 2007), which exhibited increased VD and enlarged BSCs. These alterations incorporate some of the most important changes required for C_4 photosynthesis.

Some Cleomaceae species exhibit phenotypic plasticity in leaf development and cell biology (Marshall *et al.,* 2007; Reeves *et al.,* 2018), and some species have already undergone gene duplication events (Schranz and Mitchell-Olds, 2006; Bergh *et al.,* 2014). However, modern phylogenetic methods allow us to analyze much more than just phylogeny as a statistical control. Specifically, they allow us to appreciate the evolutionary history of species, to better understand the patterns of biological diversity, to trace character traits over the evolutionary time, and to draw inferences regarding the evolution of these traits (Adams, 2014). As such, feature-based studies offer an important tool to better understand the ecological drivers of biological diversity. First, as observed in the present study, genome size was significantly correlated to variables essential for the C_4 photosynthetic mechanism, such as the $\delta^{13}C$ value, WUE*i*, A_{gross}, A_{N}, malate, aspartate, starch, proteins, venation density and palisade parenchyma (Figures 2-3). As demonstrated for Grasses (Wang *et al.*, 2009) and for *G. gynandra* – Cleomaceae (Huang *et al.*, 2021) and reported in review papers (*e.g.* Monson, 2003; Sage, 2004, 2021; Gowik and Westhoff, 2011), the genome duplication is thought to facilitate the evolution of C_4 photosynthesis from C_3 photosynthesis.

Furthermore, the second step in the development of the C_4 mechanism is characterized by increased VD (Sage 2004; Gowik and Westhoff 2011). In the present study, 15 accessions exhibited higher VD than the C_4 species *G. gynandra*; four accessions exhibited values similar to *G. gynandra*; and only two accessions exhibited values lower than *G. gynandra* (Figure 4G). Of note, the VD in *G. gynandra* was equivalent to that in other accessions of the species (6–10 $mm.mm^{-2}$) (Marshall *et al.,* 2007; Reeves *et al.,* 2018). However, in the remaining C_3 accessions studied here, VD was higher than the previously reported values in the C_3 species *C. violaceae*, *C. isomeris*, *C. hirta*, and *C. africana* as well as the C_3-C_4 intermediate species *Coalisina paradoxa* (Marshall *et al.,* 2007). Further, differences in VD observed in the present study may also be related to increase in the mechanical integrity of leaves, which may be beneficial in windy habitats or may improve water supply to leaves in dry and hot biotopes (Sage 2004). In this regard, the characteristics that may initially be an adaptation to the environment, such as the combination of shorter ID, higher VD, and larger BSC, may predispose a species to developing the C_4 photosynthetic mechanism, as evidenced in grasses (Christin *et al.*, 2013a). Likewise, in Brassicaceae plants, VD has never been reported at the level of the C_4 species, indicating that this anatomical attribute may be one of the major constraints to the evolution of C_4 traits in this family (Schlüter *et al.*, 2017). Therefore, increase in VD and enlargement of BSCs are the key alterations of foliar anatomy occurring in the C_3 context, preceding the emergence of the C_4 syndrome (Christin *et al.*, 2013a).

The third step in the development of the C_4 mechanism is marked by increase in the number of organelles in BSCs, resulting in cell enlargement. Although the activation of BSCs may be a secondary effect of increase in VD (Gowik and Westhoff, 2011), some accessions such as *T. hassleriana* (Canaã, MG, and Canoinhas, SC), *T. siliculifera*, and *T. longicarpa* (Manaus, AM) (Supplementary Figure S7) showed BSC size comparable to *G. gynandra*. The fourth step in the evolution of the C_4 mechanism is the appearance of glycine shuttling, a process in which the photorespiratory intermediate glycine is shuttled from the mesophyll tissue to the BSCs, where it is metabolized to CO_2 and serine by the action of glycine decarboxylase (Sage 2004). Plants exhibiting this trait are considered C_3-C_4 intermediates because they present intermediate characteristics between the C_3 and C_4 mechanisms.

According to a previous carbon isotope discrimination analysis, *T. siliculifera* may present a modified photosynthetic mechanism (Voznesenskaya *et al.,* 2007). However, in the present study, the *T. siliculifera* accession exhibited a $\delta^{13}C$ value equivalent to the C_3 species, although anatomical and metabolic data indicated that it may be in transition from the C_3 to C_3-C_4 metabolism. Together, these results indicate that the studied *T. siliculifera* accession warrants further investigation, at both physiological and molecular levels, to define its photosynthetic metabolism.

CONCLUSIONS

Cleomaceae is a highly diverse family in which we can observe various types of photosynthetic metabolism, in addition to morphological differences. This diversity may be related to the evolutionary history of specific clusters, as distinct groups have different genome sizes, even within the same genus. However, this conclusion must be drawn with caution, because further evidence demonstrating that other C_4 species in the Cleomaceae group consistently have large genomes is essential. In the present study, we also demonstrated the potential of Cleomaceae species in offering novel insights into the evolution of C_4 photosynthetic mechanism, particularly the initial transition from a C_3 character state, in *Tarenaya* species. The C_3 species described here exhibit increases in VD and enlargement of BSCs, which may predispose these lineages to the development of C_4 photosynthesis.

ACKNOWLEDGMENT

The authors thank the Instituto Estadual de Floresta (IEF) for the license granted (077/2017).

FUNDING

This work was supported by funding from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; grant number 424024/2018-7) and the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG; grant numbers APQ-00528-18 and CRA-RED00053-16). The authors also acknowledge the research fellowships granted by CNPq to ANN and WLA as well as the scholarship granted by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to DFP. MGMVV was supported by scholarships from CAPES (PNPD-1638006). APMW acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC-2048/1 (project ID 390686111) and by ERA-CAPS project C4BREED (WE2231/20-1).

REFERENCES

Adams, D. C. (2014). A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution (N. Y).* 68, 2675–2688. doi:10.1111/evo.12463.

Ammal, J. E. K. (1933). The chromosome number of Cleome viscosa Linn. *Curr. Sci.*, 328–328. doi:10.1016/B978-0-12-409548-9.10607-4.

Arumuganathan, K., Earle, E.D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*. 9, 208–218.

Bai, C., Alverson, W.S., Follansbee, A., Waller, D.M. (2012). New reports of nuclear DNA content for 407 vascular plant taxa from the United States. *Annals of Botany*, 110, 1623–1629. doi: 10.1093/aob/mcs222

Bayat, S., Schranz, M. E., Roalson, E. H., and Hall, J. C. (2018). Lessons from Cleomaceae, the Sister of Crucifers. *Trends Plant Sci.* 23, 808–821. doi:10.1016/j.tplants.2018.06.010.

Bergh, E. Van Den, Külahoglu, C., Bräutigam, A., Hibberd, J. M., Weber, A. P. M., Zhu, X., et al. (2014). Current Plant Biology Gene and genome duplications and the origin of C4 photosynthesis:Birth of a trait in the Cleomaceae. *Curr. Plant Biol.* 1, 2–9. doi:10.1016/j.cpb.2014.08.001.

Beric Donald Danforth, A., Mabry, M. E., Harkess, A. E., Eric Schranz, M., Conant, G. C., Edger, P. P., et al. (2020). Surprising amount of stasis in repetitive genome content across the Brassicales. *bioRxiv*, 2020.06.15.153296. doi:10.1101/2020.06.15.153296.

Bianconi, M. E., Dunning, L. T., Moreno-Villena, J. J., Osborne, C. P., and Christin, P. A. (2018). Gene duplication and dosage effects during the early emergence of C 4 photosynthesis in the grass genus Alloteropsis. *J. Exp. Bot.* 69, 1967–1980. doi:10.1093/jxb/ery029.

Bräutigam, A., and Gowik, U. (2016). Photorespiration connects C3 and C4 photosynthesis. *J. Exp. Bot.* 67, 2953–2962. doi:10.1093/jxb/erw056.

Brown, N. J., Parsley, K., and Hibberd, J. M. (2005). The future of C4 research – maize, Flaveria or Cleome? *Trends Plant Sci.* 10, 215–221. doi:10.1016/j.tplants.2005.03.003.

Caemmerer, S. Von, Ghannoum, O., Pengelly, J. J. L., and Cousins, A. B. (2014). Carbon isotope discrimination as a tool to explore C4 photosynthesis. 65, 3459–3470. doi:10.1093/jxb/eru127.

Chris Pires, J., Maureira, I. J., Givnish, T. J., Systma, K. J., Seberg, O., and Chris, J. (2006).

Phylogeny, Genome Size, and Chromosome Evolution of Asparagales. *Aliso: A Journal of Systematic and Evolutionary Botany*: 22, 287–304.

Christin, P. A., Besnard, G., Edwards, E. J., and Salamin, N. (2012a). Effect of genetic convergence on phylogenetic inference. *Mol. Phylogenet. Evol.* 62, 921–927. doi:10.1016/j.ympev.2011.12.002.

Christin, P. A., Edwards, E. J., Besnard, G., Boxall, S. F., Gregory, R., Kellogg, E. A., et al. (2012b). Adaptive evolution of C 4 photosynthesis through recurrent lateral gene transfer. *Curr. Biol.* 22, 445–449. doi:10.1016/j.cub.2012.01.054.

Christin, P. A., Osborne, C. P., Chatelet, D. S., Columbus, J. T., Besnard, G., Hodkinson, T. R., et al. (2013a). Anatomical enablers and the evolution of C4 photosynthesis in grasses. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1381–1386. doi:10.1073/pnas.1216777110.

Christin, P. A., Osborne, C. P., Chatelet, D. S., Columbus, J. T., Besnard, G., Hodkinson, T. R., et al. (2013b). Anatomical enablers and the evolution of C4 photosynthesis in grasses. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1381–1386. doi:10.1073/pnas.1216777110.

Christin, P. A., Osborne, C. P., Sage, R. F., Arakaki, M., and Edwards, E. J. (2011a). C 4 eudicots are not younger than C4 monocots. *J. Exp. Bot.* 62, 3171–3181. doi:10.1093/jxb/err041.

Christin, P. A., Sage, T. L., Edwards, E. J., Ogburn, R. M., Khoshravesh, R., and Sage, R. F. (2011b). Complex evolutionary transitions and the significance of c3-c4 intermediate forms of photosynthesis in molluginaceae. *Evolution (N. Y).* 65, 643–660. doi:10.1111/j.1558-5646.2010.01168.x.

Christin, P., Osborne, C. P., Chatelet, D. S., Columbus, J. T., Besnard, G., Hodkinson, T. R., et al. (2013c). Anatomical enablers and the evolution of C 4 photosynthesis in grasses. 110. doi:10.1073/pnas.1216777110.

Cross, J. M., Von Korff, M., Altmann, T., Bartzetko, L., Sulpice, R., Gibon, Y., et al. (2006). Variation of enzyme activities and metabolite levels in 24 arabidopsis accessions growing in carbon-limited conditions. *Plant Physiol.* 142, 1574–1588. doi:10.1104/pp.106.086629.

De Baerdemaeker, N. J. F., Hias, N., Van den Bulcke, J., Keulemans, W., and Steppe, K. (2018). The effect of polyploidization on tree hydraulic functioning. *Am. J. Bot.* 105, 161–171. doi:10.1002/ajb2.1032.

Defoort, J., Van de Peer, Y., and Carretero-Paulet, L. (2019). The evolution of gene duplicates in angiosperms and the impact of protein-protein interactions and the mechanism of duplication. *Genome Biol. Evol.* doi:10.1093/gbe/evz156.

Díez, C. M., Gaut, B. S., Meca, E., Scheinvar, E., Montes-Hernandez, S., Eguiarte, L. E., et al. (2013). Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytol.* 199, 264–276. doi:10.1111/nph.12247.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Felsenstein J. 1985*a*. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.

Felsenstein J. 1985*b*. Phylogenies and the comparative method. *American Naturalist* 125, 1–15.

Feodorova TA, Voznesenskaya E V, Edwards GE, Roalson EH. 2010. Biogeographic Patterns of Diversification and the Origins of C4 in Cleome (Cleomaceae). *Systematic Botany* **35**, 811–826.

Fernie, A. R., Roscher, A., Ratcliffe, R. G., and Kruger, N. J. (2001). Fructose 2,6-bisphosphate activates pyrophosphate: fructose-6-phosphate 1-phosphotransferase and increases triose phosphate to hexose phosphate cycling in heterotrophic cells. *Planta* 212, 250–263. doi:10.1007/s004250000386.

Fiehn, O. (2007). "Validated High Quality Automated Metabolome Analysis of Arabidopsis Thaliana Leaf Disks," in *Concepts in Plant Metabolomics* (Springer Netherlands), 1–18. doi:10.1007/978-1-4020-5608-6_1.

Flexas, J., Ortuño, M. F., Ribas-Carbo, M., Diaz-Espejo, A., Flórez-Sarasa, I. D., and Medrano, H. (2007). Mesophyll conductance to CO2 in Arabidopsis thaliana. *New Phytol.* 175, 501–511. doi:10.1111/j.1469-8137.2007.02111.x.

Flora do Brasil - Cleomaceae Bercht. & J.Presl Available at: http://floradobrasil.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/BemVindoConsultaPublicaConsultar.do?invalidatePageControlCounter=1&idsFilhosAlgas=%5B2%5D&idsFilhosFungos=%5B1%2C10%2C11%5D&lingua=&grupo=5&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=Cleomaceae+Bercht.+%26+J.Presl&formaVida=null&substrato=null&ocorreBrasil=QUALQUER&ocorrencia=OCORRE&endemismo=TODOS&origem=TODOS&regiao=QUALQUER&estado=QUALQUER&ilhaOceanica=32767&domFitogeograficos=QUALQUER&bacia=QUALQUER&vegetacao=TODOS&mostrarAte=SUBESP_VAR&opcoesBusca=TODOS_OS_NOMES&loginUsuario=Visitante&senhaUsuario=&contexto=consulta-publica [Accessed February 17, 2021].

Genty, B., Briantais, J. M., and Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochim. Biophys. Acta - Gen. Subj.* 990, 87–92. doi:10.1016/S0304-4165(89)80016-9.

Gowik, U., and Bra, A. (2011). Evolution of C 4 Photosynthesis in the Genus Flaveria : How Many and Which Genes Does It Take to Make C 4 ? 23, 2087–2105. doi:10.1105/tpc.111.086264.

Gowik, U., and Westhoff, P. (2011). The Path from C 3 to C 4 Photosynthesis 1. 155, 56–63. doi:10.1104/pp.110.165308.

Gregory, T. R. (2002). Genome size and developmental complexity. *Genetica* 115, 131–46. doi:10.1023/a:1016032400147.

Grime, J. P., and Mowforth, M. A. (1982). Variation in genome size - An ecological interpretation. *Nature* 299, 151–153. doi:10.1038/299151a0.

Gu, J., Weber, K., Klemp, E., Winters, G., Franssen, S. U., Wienpahl, I., et al. (2012). Identifying core features of adaptive metabolic mechanisms for chronic heat stress attenuation contributing to systems robustness. *Integr. Biol.* 4, 480–493. doi:10.1039/c2ib00109h.

Hall, J. C., Sytsma, K. J., and Iltis, H. H. (2002). Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. *Am. J. Bot.* 89, 1826–1842. doi:10.3732/ajb.89.11.1826.

Hao, G. Y., Lucero, M. E., Sanderson, S. C., Zacharias, E. H., and Holbrook, N. M. (2013). Polyploidy enhances the occupation of heterogeneous environments through hydraulic related trade-offs in Atriplex canescens (Chenopodiaceae). *New Phytol.* 197, 970–978. doi:10.1111/nph.12051.

Harley, P. C., Loreto, F., Marco, G. Di, Sharkey, T. D., Scalo, M., and Al, H. E. T. (1992). Theoretical Considerations when Estimating the Mesophyll Conductance to CO2 Flux by Analysis of the Response of Photosynthesis to CO21. 1429–1436.

Hatch, M. D. (1988). C 4 photosynthesis : a unique m e n a of modified biochemistry , anatomy and ultrastructure. 895, 81–106.

Hattersley, P. W. (1984). Characterization of C4 Type Leaf Anatomy in Grasses (Poaceae). Mesophyll: Bundle Sheath Area Ratios. *Ann. Bot.* 53, 163–180. doi:10.1093/oxfordjournals.aob.a086678.

Heckmann, D., Schulze, S., Denton, A., Gowik, U., Westhoff, P., Weber, A. P. M., et al. (2013). Theory Predicting C4 Photosynthesis Evolution : Modular , Individually Adaptive Steps on a Mount Fuji Fitness Landscape. *Cell* 153, 1579–1588. doi:10.1016/j.cell.2013.04.058.

Holaday, A. S., Lee, K. W., and Chollet, R. (1984). C3-C4 Intermediate species in the genus Flaveria: leaf anatomy, ultrastructure, and the effect of O2 on the CO2 compensation concentration. *Planta* 160, 25–32. doi:10.1007/BF00392462.

Huang, C.-F., Liu, W.-Y., Lu, M.-Y. J., Chen, Y.-H., Ku, M. S. B., and Li, W.-H. (2021b). Whole-Genome Duplication Facilitated the Evolution of C4 Photosynthesis in *Gynandropsis gynandra*. *Mol. Biol. Evol.* doi:10.1093/MOLBEV/MSAB200.

Iltis, H. H. (1957). Studies in the Capparidaceae. III. Evolution and Phylogeny of the Western North American Cleomoideae. *Ann. Missouri Bot. Gard.* 44, 77. doi:10.2307/2394679.

Iltis, H. H. (1959). mens above the basic number of 6 , therefore , seems to be a derived rather than and Ranmanissa have many stamens , in some species of the latter often close to.

Iltis, H. H., and Cochrane, T. S. (2014). Studies in the Cleomaceae VI : A New Genus and Sixteen New Combinations for the Flora Mesoamericana Studies in the Cleomaceae VI : A New Genus and Sixteen New Combinations for the Flora Mesoamericana. doi:10.3417/2013017.

Inda, L. A., Torrecilla, Æ. P., Catala, Æ. P., and Ruiz-zapata, T. (2008). Phylogeny of Cleome L . and its close relatives Podandrogyne Ducke and Polanisia Raf . (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. 111–126. doi:10.1007/s00606-008-0026-y.

Kadereit, G., and Freitag, H. (2011). Molecular phylogeny of Camphorosmeae (Camphorosmoideae, Chenopodiaceae): Implications for biogeography, evolution of C_4 -photosynthesis and taxonomy. *Taxon* 60, 51–78. doi:10.1002/tax.601006.

Khoshravesh, R., Stata, M., Busch, F. A., Saladié, M., Castelli, J. M., Dakin, N., et al. (2020). The Evolutionary Origin of C4 Photosynthesis in the Grass Subtribe Neurachninae 1[OPEN]. doi:10.1104/pp.19.00925.

Koteyeva, N. K., Voznesenskaya, E. V., Cousins, A. B., and Edwards, G. E. (2014). Differentiation of C4 photosynthesis along a leaf developmental gradient in two Cleome species having different forms of Kranz anatomy. *J. Exp. Bot.* 65, 3525–3541. doi:10.1093/jxb/eru042.

Koteyeva, N. K., Voznesenskaya, E. V, Roalson, E. H., and Edwards, G. E. (2011). Diversity in forms of C 4 in the genus Cleome (Cleomaceae). 269–283. doi:10.1093/aob/mcq239.

Logan, B. A., Adams, W. W., and Demmig-Adams, B. (2007). Viewpoint: Avoiding common pitfalls of chlorophyll fluorescence analysis under field conditions. *Funct. Plant Biol.* 34, 853. doi:10.1071/FP07113.

Loureiro, J., Rodriguez, E., Doležel, J., and Santos, C. (2006). Comparison of Four Nuclear Isolation Buffers for Plant DNA Flow Cytometry. *Ann. Bot.* 98, 679–689. doi:10.1093/aob/mcl141.

Maherali, H., Walden, A. E., and Husband, B. C. (2009). Genome duplication and the evolution of physiological responses to water stress. *New Phytol.* 184, 721–731. doi:10.1111/j.1469-8137.2009.02997.x.

Mallmann, J., Heckmann, D., Bräutigam, A., Lercher, M. J., Weber, A. P. M., Westhoff, P., et al. (2014). The role of photorespiration during the evolution of C4 photosynthesis in the genus Flaveria. *Elife* 2014, 1–23. doi:10.7554/eLife.02478.

Marshall, D. M., Muhaidat, R., Brown, N. J., Liu, Z., Stanley, S., Griffiths, H., et al. (2007). Cleome, a genus closely related to Arabidopsis, contains species spanning a developmental progression from C3 to C4 photosynthesis. *Plant J.* 51, 886–896. doi:10.1111/j.1365-313X.2007.03188.x.

Mckown, A. D., and Dengler, N. G. (2007). Key innovations in the evolution of Kranz anatomy and C4 vein pattern in Flaveria (Asteraceae). *Am. J. Bot.* 94, 382–389. doi:10.3732/ajb.94.3.382.

McKown, A. D., Moncalvo, J. M., and Dengler, N. G. (2005). Phylogeny of Flaveria (Asteraceae) and inference of C4 photosynthesis evolution. *Am. J. Bot.* 92, 1911–1928. doi:10.3732/ajb.92.11.1911.

Meyerson, L. A., Pyšek, P., Lučanová, M., Wigginton, S., Tran, C., and Cronin, J. T. (2020). Plant genome size influences stress tolerance of invasive and native plants via plasticity. *Ecosphere* 11. doi:10.1002/ecs2.3145.

Monson, R. K. (2003). Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. *Int. J. Plant Sci.* 164. doi:10.1086/368400.

Naganowska, B., Wolko, B., Śliwińska, E., Kaczmarek, Z. (2003). Nuclear DNA Content Variation and Species Relationships in the Genus Lupinus (Fabaceae). *Annals of Botany*. 92, 349–355. doi:10.1093/aob/mcg145.

Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., et al. (2005). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* 3, 0976–0985. doi:10.1371/journal.pbio.0030170.

Nunes-Nesi, A., Carrari, F., Gibon, Y., Sulpice, R., Lytovchenko, A., Fisahn, J., et al. (2007). Deficiency of mitochondrial fumarase activity in tomato plants impairs photosynthesis via an effect on stomatal function. *Plant J.* 50, 1093–1106. doi:10.1111/j.1365-

313X.2007.03115.x.

Nylander (2004). MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

O'Brien, T. P., Feder, N., and McCully, M. E. (1964). Polychromatic staining of plant cell walls by toluidine blue O. *Protoplasma* 59, 368–373. doi:10.1007/BF01248568.

Oa, N. D. C. W., Coate, J. E., Schlueter, J. A., Whaley, A. M., and Doyle, J. J. (2011). Comparative Evolution of Photosynthetic Genes in Response to Polyploid and. 155, 2081–2095. doi:10.1104/pp.110.169599.

Omondi, E. O., Debener, T., Linde, M., Abukutsa-Onyango, M., Dinssa, F. F., and Winkelmann, T. (2017). Mating biology, nuclear DNA content and genetic diversity in spider plant (Cleome gynandra) germplasm from various African countries. *Plant Breed.* 136, 578–589. doi:10.1111/pbr.12485.

Otto, F. (1990). DAPI Staining of Fixed Cells for High-Resolution Flow Cytometry of Nuclear DNA. *Methods Cell Biol.* 33, 105–110. doi:10.1016/S0091-679X(08)60516-6.

Otto, S. P., and Whitton, J. (2000). Polyploid Incidence and Evolution. *Annu. Rev. Genet.* 34, 401–437. doi:10.1146/annurev.genet.34.1.401.

Patchell, M. J., Roalson, E. H., and Hall, J. C. (2014). Resolved phylogeny of Cleomaceae based on all three genomes. 63, 315–328.

Paterson, A. H., Freeling, M., Tang, H., and Wang, X. (2010). Insights from the Comparison of Plant Genome Sequences. *Annu. Rev. Plant Biol.* 61, 349–372. doi:10.1146/annurev-arplant-042809-112235.

Pons, T. L., Flexas, J., Caemmerer, S. Von, Evans, J. R., Genty, B., Ribas-carbo, M., et al. (2009). Estimating mesophyll conductance to CO 2 : methodology , potential errors , and recommendations. 60, 2217–2234. doi:10.1093/jxb/erp081.

Porra, R. J., Thompson, W. A., and Kriedemann, P. E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *BBA - Bioenerg.* 975, 384–394. doi:10.1016/S0005-2728(89)80347-0.

Posada, D., and Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53, 793–808. doi:10.1080/10635150490522304.

Praça-Fontes, M. M., Carvalho, C. R., Clarindo, W. R., and Cruz, C. D. (2011). Revisiting the

DNA C-values of the genome size-standards used in plant flow cytometry to choose the "best primary standards." *Plant Cell Rep.* 30, 1183–1191. doi:10.1007/s00299-011-1026-x.

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., et al. (2019a). Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biol.* 20, 38. doi:10.1186/s13059-019-1650-2.

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., et al. (2019b). Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biol.* 20, 38. doi:10.1186/s13059-019-1650-2.

Reeves, G., Singh, P., Rossberg, T. A., Sogbohossou, E. O. D., Schranz, M. E., and Hibberd, J. M. (2018). Natural variation within a species for traits underpinning C 4 photosynthesis. *Plant Physiol.* 177, 504–512. doi:10.1104/pp.18.00168.

Roddy, A. B., Théroux-Rancourt, G., Abbo, T., Benedetti, J. W., Brodersen, C. R., Castro, M., et al. (2020). The scaling of genome size and cell size limits maximum rates of photosynthesis with implications for ecological strategies. *Int. J. Plant Sci.* 181, 75–87. doi:10.1086/706186.

Rodeghiero, M., Niinemets, Ü., and Cescatti, A. (2007). Major diffusion leaks of clamp-on leaf cuvettes still unaccounted: How erroneous are the estimates of Farquhar et al. model parameters? *Plant, Cell Environ.* 30, 1006–1022. doi:10.1111/j.1365-3040.2007.001689.x.

Sage, R. F. (2001). Environmental and evolutionary preconditions for the origin and diversification of the C4 photosynthetic syndrome. *Plant Biol.* 3, 202–213. doi:10.1055/s-2001-15206.

Sage, R. F. (2004). The evolution of C4 photosynthesis. *New Phytol.* 161, 341–370. doi:10.1111/j.1469-8137.2004.00974.x.

Sage, R. F., Christin, P., and Edwards, E. J. (2011). The C4 plant lineages of planet Earth. 62, 3155–3169. doi:10.1093/jxb/err048.

Sage, R. F., Khoshravesh, R., and Sage, T. L. (2014). From proto-Kranz to C4 Kranz: building the bridge to C 4 photosynthesis. *J. Exp. Bot.* 65, 3341–3356. doi:10.1093/jxb/eru180.

Sanchez-Acebo, L. (2005). A Phylogenetic Study of the New World Cleome (Brassicaceae, Cleomoideae) on JSTOR. *Ann. Missouri Bot. Gard.* 92, 179–201. Available at: https://www.jstor.org/stable/3298513?seq=1#metadata_info_tab_contents [Accessed February 27, 2021].

Sattler, M. C., Carvalho, C. R., and Clarindo, W. R. (2016). The polyploidy and its key role in plant breeding. *Planta* 243, 281–296. doi:10.1007/s00425-015-2450-x.

Schranz, M. E., and Mitchell-olds, T. (2006). Independent Ancient Polyploidy Events in the Sister Families Brassicaceae and Cleomaceae. 18, 1152–1165. doi:10.1105/tpc.106.041111.1.

Schuler, M. L., Mantegazza, O., and Weber, A. P. M. (2016). Engineering C 4 photosynthesis into C 3 chassis in the synthetic biology age. 51–65. doi:10.1111/tpj.13155.

Shim, S.-H., Lee, S.-K., Lee, D.-W., Brilhaus, D., Wu, G., Ko, S., et al. (2020). Loss of Function of Rice Plastidic Glycolate/Glycerate Translocator 1 Impairs Photorespiration and Plant Growth. *Front. Plant Sci.* 10, 1726. doi:10.3389/fpls.2019.01726.

Silva Artur, M. A., Costa, M.-C. D., Farrant, J. M., and Hilhorst, H. W. M. (2019). Genome-level responses to the environment: plant desiccation tolerance. *Emerg. Top. Life Sci.* 3, 153–163. doi:10.1042/etls20180139.

Simonin, K. A., and Roddy, A. B. (2018). Genome downsizing, physiological novelty, and the global dominance of flowering plants. *PLOS Biol.* 16, e2003706. doi:10.1371/journal.pbio.2003706.

Stevens, P. F. Angiosperm Phylogeny Website. Version 14, July 2017 [and more or less continuously updated since]. Available at: http://www.mobot.org/mobot/research/apweb/orders/brassicalesweb.htm [Accessed February 27, 2021].

Subramania, D., and Susheela, G. Cytotaxonomical Studies of South Indian Capparidaceae.

Tuler, A. C., Carrijo, T. T., Peixoto, A. L., Garbin, M. L., da Silva Ferreira, M. F., Carvalho, C. R., et al. (2019). Diversification and geographical distribution of Psidium (Myrtaceae) species with distinct ploidy levels. *Trees* 33, 1101–1110. doi:10.1007/s00468-019-01845-2.

U, S., and AP, Weber. (2016). The Road to C4 Photosynthesis: Evolution of a Complex Trait via Intermediary States. *Plant Cell Physiol.* 57. doi:10.1093/PCP/PCW009.

Von Caemmerer, S., and Furbank, R. T. (2003). The C4 pathway: An efficient CO2 pump. *Photosynth. Res.* 77, 191–207. doi:10.1023/A:1025830019591.

Voznesenskaya, E. V., Koteyeva, N. K., Chuong, S. D. X., Ivanova, A. N., Barroca, J., Craven, L. A., et al. (2007). Physiological, anatomical and biochemical characterisation of photosynthetic types in genus Cleome (Cleomaceae). *Funct. Plant Biol.* 34, 247. doi:10.1071/FP06287.

Voznesenskaya, E. V., Koteyeva, N. K., Cousins, A., and Edwards, G. E. (2018). Diversity in structure and forms of carbon assimilation in photosynthetic organs in Cleome (Cleomaceae). *Funct. Plant Biol.* 45, 983–999. doi:10.1071/FP17323.

Vyas, P., Bisht, M. S., Miyazawa, S.-I., Yano, S., Noguchi, K., Terashima, I., et al. (2007). Effects of polyploidy on photosynthetic properties and anatomy in leaves of Phlox drummondii. *Funct. Plant Biol.* 34, 673. doi:10.1071/FP07020.

Walker, B. J., and Cousins, A. B. (2013). Influence of temperature on measurements of the CO 2 compensation point : differences between the Laisk and O 2 - exchange methods. 64, 1893–1905. doi:10.1093/jxb/ert058.

Wang, L., Czedik-Eysenberg, A., Mertz, R. A., Si, Y., Tohge, T., Nunes-Nesi, A., et al. (2014). Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nat. Biotechnol.* 32, 1158–1165. doi:10.1038/nbt.3019.

Wang, P., Moore, B. M., Panchy, N. L., Meng, F., Lehti-Shiu, M. D., and Shiu, S. H. (2018). Factors influencing gene family size variation among related species in a plant family, solanaceae. *Genome Biol. Evol.* 10, 2596–2613. doi:10.1093/gbe/evy193.

Wang, X., Gowik, U., Tang, H., Bowers, J. E., Westhoff, P., and Paterson, A. H. (2009). Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. *Genome Biol.* 10, R68. doi:10.1186/gb-2009-10-6-r68.

Warner, D. A., and Edwards, G. E. (1993). Effects of polyploidy on photosynthesis. *Photosynth. Res.* 35, 135–147. doi:10.1007/BF00014744.

White, T. J., T. D. Bruns, S. B. Lee, and J. W. Taylor. (1990). Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. *Molecula.* doi:10.1016/b978-0-12-372180-8.50042-1

Williams, B. P., Aubry, S., and Hibberd, J. M. (2012). Molecular evolution of genes recruited into C4 photosynthesis. *Trends Plant Sci.* 17, 213–220. doi:10.1016/j.tplants.2012.01.008.

Yang, H., Yu, Q., Sheng, W. P., Li, S. G., and Tian, J. (2017). Determination of leaf carbon isotope discrimination in C4 plants under variable N and water supply. *Sci. Rep.* 7. doi:10.1038/s41598-017-00498-w.

Yin, X., Struik, P. C., Romero, P., Harbinson, J., Evers, J. B., Van Der Putten, P. E. L., et al. (2009). Using combined measurements of gas exchange and chlorophyll fluorescence to estimate parameters of a biochemical C3 photosynthesis model: A critical appraisal and a new integrated approach applied to leaves in a wheat (Triticum aestivum) canopy. *Plant, Cell Environ.* 32, 448–464. doi:10.1111/j.1365-3040.2009.01934.x.

Zhang, J., Wang, M., Guo, Z., Guan, Y., Guo, Y., and Yan, X. (2019). Variation in ploidy level and genome size of Cynodon dactylon (L.) Pers. along a latitudinal gradient. *Folia Geobot.* 54, 267–278. doi:10.1007/s12224-019-09359-y.

Zsögön, A., Alves Negrini, A. C., Peres, L. E. P., Nguyen, H. T., and Ball, M. C. (2015). A mutation that eliminates bundle sheath extensions reduces leaf hydraulic conductance, stomatal conductance and assimilation rates in tomato (*Solanum lycopersicum*). *New Phytol.* 205, 618–626. doi:10.1111/nph.13084.

Table S1. List of ITS sequences indicating species and its respective number of identification – accession number at NCBI – used for molecular phylogeny analyses.

Species	**NCBI ID**
Arabidopsis thaliana	MG886682
Areocleome oxalidea	HM044247
Arivela cleomoides	KF923166
Arivela microaustralica	HM044245
Arivela uncifera	HM044249
Arivela viscosa	KT588819
Brassica macrocarpa	GQ268076
Brassica oleracea	GQ891871
Brassica rapa	GQ891874
Cleome Africana	HM044222
Cleome amblyocarpa	KF850548
Cleome Arabica	KF805109
Cleome breyeri	HM044258
Cleome coluteoides	HM044224
Cleome densifolia	KF923168
Cleome foliosa	KF923173
Cleome houtteana	LC144536
Cleome kalachariensis	HM044277
Cleome khorassanica	HM044230
Cleome linearifolia	HM044278
Cleome luederitziana	HM044256
Cleome melanosperma	HM044284
Cleome monochrome	HM044267
Cleome moritziana	DQ455794
Cleome ornithopodioides	KF923178
Cleome ramosissima	KF850556
Cleome rotundifolia	KF923182
Cleome schweinfurthii	HM044273
Cleome stenophylla	DQ455814
Cleome strigosa	KF923184
Cleome stylosa	DQ455812
Cleome torticarpa	DQ455810
Cleome turkmena	HM044231
Cleome violacea	KF923185
Cleomella goodrichii	KF217186
Cleomella hillmanii	KF217187
Cleomella lutea	HM044235
Cleomella obtusifolia	KF217196
Cleomella plocasperma	HM044234
Cleoserrata paludosa	MF287989
Cleoserrata speciosa	KT588831
Coalisina angustifolia	HM044251
Coalisina diandra	KF923169
Coalisina paradoxa	HM044257
Corynandra aspera	KT588829
Corynandra chelidonii	KT588822
Corynandra feline	KT588830
Corynandra simplicifolia	KT588824
Dactylaena microphylla	HM044279
Dactylaena sp.	MN879423
Dactylaena sp.	MN879424
Gynandropsis gynandra	MH188313
Gynandropsis gynandra	MH188314
Kersia kalachariensis	HM044277
Melidiscus giganteus	HM044283
Peritoma arborea	KF217218
Peritoma multicaulis	KF217228

Peritoma platycarpa	KF217230
Peritoma serrulata	KF217232
Physostemon lancolatum	KF923175
Podandrogyne chiriquensis	HM044281
Podandrogyne jamesonii	HM044282
Podandrogyne macrophylla	DQ455815
Polanisia dodecandra	DQ455816
Rorida droserifolia	HM044229
Rorida fimbriata	HM044227
Rorida quinquenervia	HM044228
Sieruela allamanii	HM044270
Sieruela briquetii	KF923165
Sieruela elegantissima	HM044272
Sieruela hirta	KR734333
Sieruela iberidella	KF923174
Sieruela macrophylla	HM044262
Sieruela maculate	HM044263
Sieruela monophylla	KT588821
Sieruela oxyphylla	KF923179
Sieruela rutidosperma	DQ455802
Sieruela schimperi	HM044273
Sieruela stricta	HM044259
Sieruela usambarica	HM044274
Tarenaya aculeata	MN879435
*Tarenaya aculeata*_TA*	MT090703
Tarenaya afrospina	HM044290
Tarenaya afrospina	MN879436
Tarenaya afrospina	MN879437
Tarenaya atropurpurea	MN879438
Tarenaya bicolor	MN906009
Tarenaya boliviensis	DQ455785
Tarenaya chapalaensis	DQ455800
Tarenaya curvispina	MN879443
Tarenaya diffusa	KF923170
Tarenaya domingensis	KF923171
Tarenaya domingensis	MN879444
Tarenaya domingensis	MN879445
Tarenaya guaratinica	MN879446
Tarenaya hassleriana	DQ455791
*Tarenaya hassleriana*_THC*	MT090700
*Tarenaya hassleriana*_THCS*	MT090716
*Tarenaya hassleriana*_THDM*	MT090714
*Tarenaya hassleriana*_THJ*	MT090704
*Tarenaya hassleriana*_THP*	MT090719
*Tarenaya hassleriana*_THS*	MT090709
*Tarenaya hassleriana*_THV*	MT090705
Tarenaya horrida	MN879447
Tarenaya houtteana	LC144536
Tarenaya lilloi	MN879449
Tarenaya longicarpa _TAM*	MT090707
Tarenaya longicarpa _TARC*	MT090712
Tarenaya longicarpa _TC*	MT090719
Tarenaya longicarpa _TIB*	MT090713
*Tarenaya longicarpa*_TAF*	MT090715
*Tarenaya longicarpa*_TL*	MT090718
*Tarenaya microcarpa*_TM*	MT090701
Tarenaya parviflora	MN879456
*Tarenaya parviflora*_TP*	MT090710

Tarenaya pernambucensis	MN879458
Tarenaya pernambucensis	MN879457
Tarenaya rosea	MN879459
*Tarenaya rosea*_TR*	MT090702
Tarenaya siliculifera	HM044286
*Tarenaya siliculifera*_TSI*	MT090717
Tarenaya spinosa	HM044296
*Tarenaya spinosa*_TS*	MT090706
Tarenaya titubans	DQ455813
Tarenaya torticarpa	DQ455810
Tarenaya trachycarpa	HM044297
Tarenaya tucumanensis	MN879465
Tarenaya tucumanensis	DQ455811
Tarenaya virens	MN879466
Tarenaya werdermannii	MN879467

*ITS sequences produced in this study.

Table S2. Relative metabolite content in leaves of Cleomaceae species sampled at middle of the day. Most species are five months old, except for TA, TM, TD and GG, that are two months old. Data are presented as means ± SE. Letters indicate significant groupings according to Tukey's Test, n=5.

	CP	GG	TA	TAF	TAM	TARC	TC	TD
Aspartate	213 ± 30a	474 ± 170a	364 ± 126a	196 ± 31a	169 ± 22a	394 ± 133a	331 ± 93a	317 ± 72a
Myoinositol	3280 ± 362b	1193 ± 262b	2939 ± 895b	6014 ± 1478b	5324 ± 671b	4575 ± 624b	2962 ± 461b	7010 ± 1149b
Leucine	6 ± 1.7a	12 ± 3.3a	14 ± 3.2a	20 ± 8.8a	11 ± 2.7a	22 ± 3.8a	12 ± 4.1a	26 ± 7.9a
Methionine	14 ± 4.4ab	5 ± 2.1b	4 ± 2.0b	21 ± 5.7ab	15 ± 6.0ab	26 ± 6.4ab	26 ± 10ab	5 ± 1.5b
Putrescine	13 ± 3.6a	10 ± 2.1a	67 ± 21.0a	18 ± 4.2a	33 ± 6.8a	67 ± 18.6a	19 ± 3.2a	94 ± 21.5a
Proline	10 ± 1.5a	39 ± 16.4a	15 ± 5.5a	17 ± 1.8a	12 ± 2.3a	34 ± 6.9a	30 ± 3.8a	18 ± 2.5a
Valine	144 ± 16.6a	100 ± 20.8a	103 ± 19.1a	116 ± 12.1a	107 ± 11.8a	144 ± 14.4a	204 ± 66a	189 ± 39.7a
Succinate	70 ± 6.6b	77 ± 27.9b	64 ± 21.0b	90 ± 11.3b	57 ± 5.2b	67 ± 7.6b	97 ± 22.3b	64 ± 7.1b
Phenylalanine	120 ± 15.5a	66 ± 16.4a	131 ± 43.9a	210 ± 50.8a	82 ± 10.9a	152 ± 15.7a	141 ± 27.9a	114 ± 24.6a
Isocitrate	14 ± 3.6a	24 ± 7.5a	14 ± 1.4a	7 ± 2.3a	8 ± 1.6a	5 ± 2.0a	7 ± 1.7a	7 ± 2.7a
Glutamate	642 ± 63.3b	870 ± 103.4ab	458 ± 78.3b	611 ± 33.0b	547 ± 18.3b	791 ± 120.0b	838 ± 201.4ab	616 ± 81.2b
Glycerate	126 ± 32.9bc	64 ± 7.8bc	103 ± 22.4bc	247 ± 13.5ab	108 ± 8.6bc	129 ± 18.7bc	235 ± 7.1abc	194 ± 31.3abc
Fructose	5637 ± 860.4b	1497 ± 109.8b	3113 ± 405.5b	6439 ± 1025.0b	6500 ± 1289.0b	10386 ± 873.5b	9093 ± 1044.7b	5449 ± 1085.6b
Asparagine	1 ± 0.1a	10 ± 2.6a	3 ± 0.7a	3 ± 1.0a	1 ± 0.2a	7 ± 1.8a	5 ± 1.9a	2 ± 0.3a
Glucose	4804 ± 628.8b	1742 ± 443.2b	3555 ± 788.8b	6274 ± 1739.9b	5411 ± 1055.8b	7814 ± 652.8b	6181 ± 111.1b	8176 ± 1201.6b
Sinapinate	10 ± 3.9a	3 ± 1.4a	21 ± 5.3a	14 ± 5.3a	5 ± 1.0a	22 ± 10.6a	7 ± 1.1a	12 ± 3.1a
Malonate	6 ± 0.9b	2 ± 0.4b	7 ± 1.6b	11 ± 2.8b	6 ± 0.7b	6 ± 1.2b	6 ± 1.4b	7 ± 0.4b
Gluconate	136 ± 53.7a	42 ± 15.2a	15 ± 3.6a	42 ± 14.2a	51 ± 21.1a	52 ± 23.2a	30 ± 4.7a	39 ± 10.3a
Glycine	366 ± 157.2a	113 ± 34.0a	435 ± 111.2a	462 ± 124.5a	102 ± 10.4a	458 ± 160.9a	800 ± 340.4a	207 ± 43.4a
5-oxoproline	1412 ± 169.6a	1462 ± 356.8a	1731 ± 489.9a	1331 ± 347.3a	1261 ± 228.3a	2060 ± 484.0a	1583 ± 464.6a	3406 ± 830.4a
Gaba	24 ± 3.4a	20 ± 1.3a	30 ± 6.0a	20 ± 0.8a	26 ± 2.4a	36 ± 9.3a	11 ± 2.7a	46 ± 9.6a
Tyrosine	11 ± 2.5a	14 ± 2.4a	13 ± 3.2a	116 ± 67.8a	18 ± 4.5a	43 ± 8.6a	32 ± 10.6a	20 ± 4.1a
alpha-Alanine	359 ± 29.5ab	2271 ± 472.8ab	358 ± 66.4ab	259 ± 13.3ab	511 ± 72.9ab	804 ± 168.5ab	908 ± 270.3ab	581 ± 60.5ab
Glycerol	299 ± 26.4c	203 ± 42.2c	318 ± 50.5c	524 ± 85.8ab	298 ± 35.3c	365 ± 20.4bc	347 ± 53.9bc	425 ± 58.5bc
Malate	1008 ± 245.2b	1113 ± 399.0b	1509 ± 69.4b	3098 ± 1089.4b	1370 ± 275.7b	967 ± 220.7b	1350 ± 213.3b	1616 ± 238.8b
Threonate	423 ± 40.5b	323 ± 80.0b	284 ± 98.5b	698 ± 221.1b	350 ± 34.6b	373 ± 35.7b	376 ± 57.3b	253 ± 24.3b
Isoleucine	23 ± 4.4a	15 ± 3.6a	45 ± 12.2a	111 ± 47.2a	33 ± 5.3a	47 ± 6.4a	114 ± 57.5a	53 ± 9.5a
Fumarate	19 ± 3.6a	54 ± 22.2a	25 ± 4.1a	93 ± 43.5a	18 ± 4.1a	16 ± 1.2a	19 ± 3.5a	23 ± 1.7a
Sucrose	21285 ± 1022.9b	17052 ± 2427.5b	20219 ± 2838.0b	26102 ± 4254.4b	20662 ± 2540.2b	27077 ± 1188.0b	24464 ± 3215.3b	28833 ± 3095.3b
Citrate+Isocitrate	483 ± 92.7a	2775 ± 730.2a	1489 ± 146.8a	1535 ± 195.5a	384 ± 94.6a	287 ± 80.5a	379 ± 91.5a	1530 ± 127.2a
Threonine	91 ± 7.8a	90 ± 13.6a	113 ± 25.4a	156 ± 56.3a	134 ± 30.4a	128 ± 22.6a	88 ± 23.2a	224 ± 61.6a
Glutamine	74 ± 23.4a	327 ± 97.8a	188 ± 57.3a	63 ± 13.8a	65 ± 13.9a	543 ± 149.0a	307 ± 71.3a	181 ± 54.6a
Serine	262 ± 86.2a	194 ± 45.0a	251 ± 61.9a	231 ± 35.4a	365 ± 120.7a	975 ± 145.1a	1023 ± 127.1a	534 ± 113.3a

Continued...

	THC	THCS	THDM	THJ	THP	THS	THV
Aspartate	355 ± 75a	190 ± 24a	192 ± 53a	139 ± 41a	717 ± 96a	93 ± 19a	668 ± 92a
Myoinositol	2760 ± 576b	2984 ± 398b	2321 ± 401b	2020 ± 651b	2519 ± 342b	3342 ± 560b	3139 ± 493b
Leucine	20 ± 7.1a	19 ± 6.1a	32 ± 7.0a	8 ± 1.5a	22 ± 2.4a	11 ± 1.4a	44 ± 5.7a
Methionine	45 ± 8.8a	19 ± 4.1ab	25 ± 9.8ab	14 ± 5.3ab	40 ± 10.1ab	3 ± 0.4b	23 ± 6.0ab
Putrescine	38 ± 9.9a	4 ± 0.8a	18 ± 7.3a	29 ± 3.5a	13 ± 1.3a	3 ±1.1a	23 ± 8.1a
Proline	29 ± 6.6a	6 ± 0.5a	5 ± 0.8a	6 ± 0.9a	33 ± 5.8a	10 ± 2.2a	42 ± 7.4a
Valine	157 ± 42.1a	117 ± 9.2a	143 ± 23.7a	137 ± 4.9a	254 ± 74.4a	91 ± 7.3a	267 ± 41.7a
Succinate	83 ± 18.4b	51 ± 6.2b	54 ± 6.3b	62 ± 6.0b	170 ± 43.0b	40 ± 3.8b	103 ± 20.7b
Phenylalanine	177 ± 39.1a	98 ± 8.4a	117 ± 20.4a	111 ± 9.8a	196 ± 38.5a	74 ± 11.6a	174 ± 7.9a
Isocitrate	12 ± 2.4a	15 ± 1.6a	17 ± 3.3a	8 ± 0.6a	21 ± 5.4a	11 ± 2.2a	11 ± 3.2a
Glutamate	613 ± 65.4ab	547 ± 44.7b	396 ± 63.3b	429 ± 68.5b	2995 ± 393.2a	266 ± 29.5b	2085 ± 422.0ab
Glycerate	155 ± 8.7bc	54 ± 5.1c	61 ± 6.5bc	96 ± 8.6bc	138 ± 11abc	54 ± 7.7c	203 ± 17.3abc
Fructose	6889 ± 802.2b	1239 ± 802.2b	3002 ± 208.9b	7260 ± 1033.1b	8293 ± 383.6b	3784 ± 1760.0b	7388 ± 1178.6b
Asparagine	4 ± 0.8a	2 ± 0.5a	2 ± 0.7a	2 ± 0.3a	12 ± 1.3a	1 ± 0.2a	10 ± 1.0a
Glucose	5513 ± 804.6b	1152 ± 181.5b	2439 ± 789.0b	6046 ± 611.6b	4182 ± 307.3b	4583 ± 511.8b	5133 ± 1029.5b
Sinapinate	6 ± 1.1a	15 ± 2.9a	15 ± 3.4a	5 ± 1.3a	4 ± 0.9a	6 ± 0.9a	8 ± 1.6a
Malonate	8 ± 1.5b	8 ± 0.9b	8 ± 1.5b	5 ± 1.2b	11 ± 2.0b	10 ± 0.8b	7 ± 1.1b
Gluconate	32 ± 4.1a	25 ± 4.6a	63 ± 17.1a	141 ± 9.8a	44 ± 12.2a	94 ± 29.0a	96 ± 37.3a
Glycine	1045 ± 393.7a	71 ± 4.8a	176 ± 73.6a	257 ± 27.0a	1219 ± 486.9a	62 ± 6.5a	417 ± 48.6a
5-oxoproline	3809 ± 685.2a	627 ± 90.5a	736 ± 247.8a	924 ± 208.7a	2599 ± 369.2a	431 ± 82.3a	2657 ± 505.2a
Gaba	51 ± 13.2a	34 ± 4.9a	36 ± 5.6a	32 ± 6.8a	55 ± 1.9a	25 ± 0.7a	65 ± 5.9a
Tyrosine	48 ± 13.6a	22 ± 6.6a	54 ± 10.9a	29 ± 8.1a	48 ± 12.9a	15 ± 3.9a	45 ± 11.4a
alpha-Alanine	758 ± 140.7ab	288 ± 26.8b	494 ± 243.3ab	469 ± 17.2ab	2252 ± 787.5a	257 ± 51.9b	920 ± 142.9ab
Glycerol	330 ± 54.8ab	264 ± 19.0c	305 ± 25.1c	312 ± 13.6c	338 ± 24.8bc	246 ± 12.0c	353 ± 29.5bc
Malate	549 ± 96.4b	573 ± 127.4b	492 ± 46.4b	722 ± 181.6b	1498 ± 352.1b	647 ± 214.1b	718 ± 191.2b
Threonate	251 ± 35.9b	381 ± 48.4b	321 ± 35.3b	428 ± 61.3b	394 ± 71.9b	416 ± 75.1b	446 ± 47.0b
Isoleucine	57 ± 20.3a	33 ± 6.6a	43 ± 14.4a	23 ± 2.7a	69 ± 24.8a	20 ± 2.6a	97 ± 26.8a
Fumarate	19 ± 3.1a	9 ± 1.1a	11 ± 2.3a	12 ± 1.0a	50 ± 16.5a	6 ± 1.1a	18 ± 2.8a
Sucrose	24276 ± 3792.5b	16564 ± 1210.4b	18553 ± 651.2b	15076 ± 592.4b	32268 ± 2515.7b	17684 ± 460.7b	27542 ± 2387.3b
Citrate+Isocitrate	1312 ± 160.1a	2081 ± 597.1a	1105 ± 104.9a	300 ± 42.0a	1308 ± 134.7a	611 ± 201.0a	1060 ± 243.9a
Threonine	130 ± 18.9a	83 ± 11.4a	78 ± 8.8a	56 ± 7.4a	157 ± 16.4a	52 ± 3.5a	159 ± 29.7a
Glutamine	157 ± 22.4a	25 ± 6.9a	11 ± 2.9a	45 ± 17.2a	485 ± 147.5a	9 ± 1.5a	373 ± 88.5a
Serine	933 ± 115.0a	192 ± 67.5a	147 ± 57.6a	256 ± 56.4a	2113 ± 745.2a	39 ± 3.1a	1170 ± 293.8a

Continued...

	TIB	TL	TM	TP	TR	TS	TSI
Aspartate	752 ± 31a	147 ± 19a	363 ± 43a	224 ± 76a	370 ± 95a	88 ± 14a	272 ± 17a
Myoinositol	4165 ± 846b	3559 ± 635b	3393 ± 1005b	3679 ± 916b	4062 ± 1018b	3313 ± 323b	3608 ± 55a
Leucine	20 ± 3.6a	8 ± 2.1a	7.8 ± 2.0a	5 ± 0.6a	13 ± 2.1a	6 ± 0.8a	38 ± 4.2a
Methionine	15 ± 4.5ab	13 ± 3.4ab	3 ± 0.7b	9 ± 2.7ab	21 ± 5.2ab	4 ± 0.6ab	11 ± 0.5ab
Putrescine	57 ± 21.2a	21 ± 4.2a	4 ± 1.4a	26 ± 8.6a	50 ± 15.1a	23 ± 7.2a	22 ± 5.0a
Proline	76 ± 24.8a	23 ± 4.4a	55 ± 18.4a	6.9 ± 0.8a	21 ± 6.5a	5 ± 0.9a	53 ± 1.5a
Valine	187 ± 47.7a	158 ± 46.1a	132 ± 16.6a	72 ± 8.7a	122 ± 9.4a	100 ± 13.3a	250 ± 1.2a
Succinate	92 ± 19.4b	123 ± 26.6b	64 ± 14.9b	63 ± 10.7b	63 ± 7.0b	36 ± 6.3b	384 ± 3.1a
Phenylalanine	156 ± 37.1a	123 ± 20.4a	124 ± 30.8a	55 ± 5.3a	107 ± 17.2a	63 ± 9.4a	202 ± 20.8a
Isocitrate	12 ± 4.7a	3 ± 2.1a	19 ± 5.3a	11 ± 2.4a	6 ± 1.4a	6 ± 2.4a	23 ± 3.4a
Glutamate	1682 ± 649ab	593 ± 73.7b	887 ± 108.7b	296 ± 48.9b	598 ± 70.5b	266 ± 32.3b	1765 ± 78ab
Glycerate	128 ± 26.2abc	173 ± 13.7abc	117 ± 26.4bc	58 ± 1.7bc	157 ± 10.0bc	59 ± 8.1bc	644 ± 13.1a
Fructose	6774 ± 1169.6b	7721 ± 1661.2b	6218 ± 748.2b	5673 ± 1255.5b	5616 ± 476.8b	5255 ± 183.1b	46737 ± 623.4a
Asparagine	4 ± 0.7a	2 ± 0.6a	4 ± 0.6a	1 ± 0.3a	3 ± 1.3a	1 ± 0.4a	10 ± 0.3a
Glucose	5147 ± 1339.1b	7876 ± 986.2b	8288 ± 679.9b	4027 ± 661.2b	4109 ± 385.1b	4235 ± 272.8b	40997 ± 423.9a
Sinapinate	5 ± 1.8a	5 ± 1.0a	4 ± 0.5a	3 ± 0.2a	4 ± 0.7a	3 ± 0.3a	22 ± 0.4a
Malonate	8 ± 0.8b	5 ± 1.3b	6 ± 1.1b	5 ± 0.8b	8 ± 1.3b	4 ± 0.3b	61 ± 6.3b
Gluconate	26 ± 5.5a	29 ± 5.8a	29 ± 8.7a	74 ± 23.4a	50 ± 16.7a	42 ± 19.4a	94 ± 6.3a
Glycine	680 ± 31.6a	843 ± 316.1a	520 ± 234.2a	76 ± 11.8a	381 ± 149.0a	72 ± 8.3a	392 ± 58.1a
5-oxoproline	2660 ± 235.7a	1461 ± 453.1a	2057 ± 248.3a	780 ± 234.4a	1976 ± 600.9a	719 ± 188.9a	1730 ± 20.5a
Gaba	17 ± 1.8a	33 ± 3.9a	30 ± 4.4a	17 ± 2.6a	40 ± 3.6a	16 ± 1.0a	53 ± 0.4a
Tyrosine	48 ± 20.0a	16 ± 3.6a	22 ± 4.3a	9 ± 1.5a	26 ± 7.1a	13 ± 4.2a	42 ± 0.2a
alpha-Alanine	990 ± 342.5ab	872 ± 223.7ab	824 ± 81.6ab	825 ± 333.3ab	733 ± 187.7ab	276 ± 35.5b	1130 ± 27.7ab
Glycerol	277 ± 42.9bc	274 ± 13.2c	327 ± 24.5c	245 ± 24.7c	351 ± 18.2bc	236 ± 40.5c	1316 ± 10.0a
Malate	1883 ± 591.0b	1659 ± 140.6b	1887 ± 397.7b	899 ± 121.1b	1160 ± 228.5b	828 ± 217.3b	7838 ± 71.5a
Threonate	405 ± 53.4b	391 ± 74.7b	272 ± 58.1b	492 ± 60.6b	453 ± 37.2b	290 ± 59.5b	2626 ± 17.8a
Isoleucine	79 ± 23.7a	80 ± 26.4a	44 ± 11.0a	16 ± 2.2a	55 ± 18.6a	17 ± 2.3a	63 ± 0.3a
Fumarate	41 ± 16.9a	27 ± 5.6a	27 ± 4.9a	20 ± 6.9a	26 ± 7.7a	12 ± 3.6a	89 ± 1.0a
Sucrose	27351 ± 4156.5b	22400 ± 1757.7b	24093 ± 3808.8b	16067 ± 2358.8b	22356 ± 2898.6b	16338 ± 1711.0b	133716 ± 110.0a
Citrate+Isocitrate	741 ± 127.1a	386 ± 120.3a	2157 ± 303.3a	259 ± 39.8a	298 ± 36.7a	160 ± 43.3a	783 ± 1.6a
Threonine	179 ± 55.5a	129 ± 48.1a	162 ± 20.2a	55 ± 6.7a	137 ± 29.8a	58 ± 9.7a	150 ± 0.6a
Glutamine	237 ± 109.8a	422 ± 49.1a	304 ± 49.0a	26 ± 4.6a	313 ± 150.0a	22 ± 5.7a	17 ± 0.6a
Serine	687 ± 222.0a	663 ± 59.3a	361 ± 50.6a	170 ± 32.7a	379 ± 87.0a	118 ± 49.6a	83 ± 0.1a

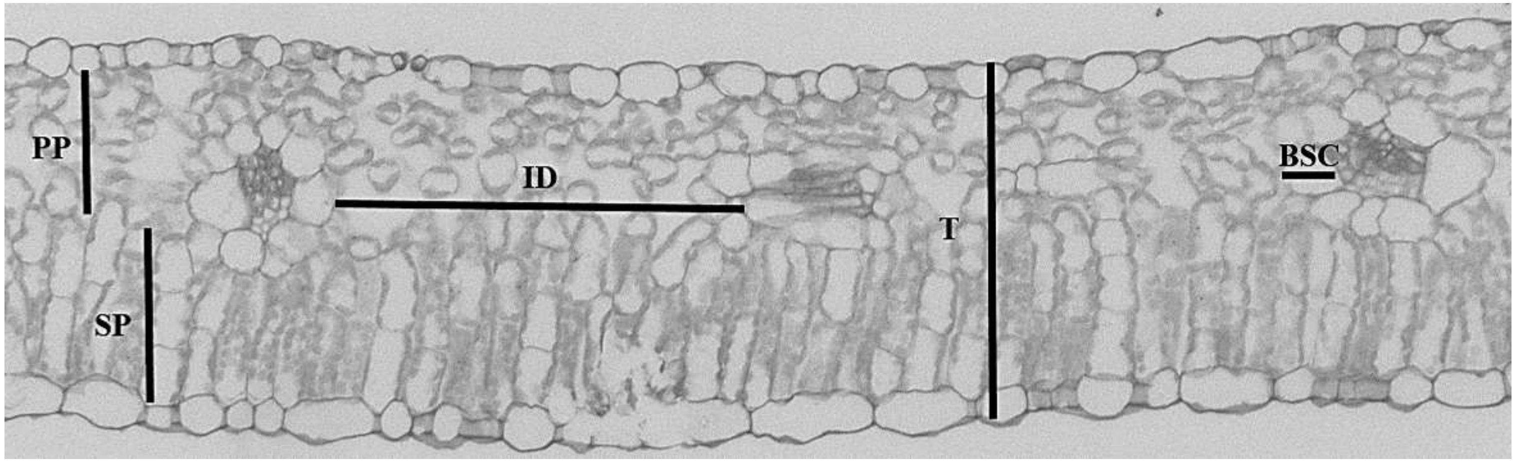


Figure S1. Scheme of how measurements were performed in the anatomical cross-section. PP: palisade parenchyma; SP: spongy parenchyma; T: leaf thickness; BSC: bundle sheath cell width; ID: internerval distance.

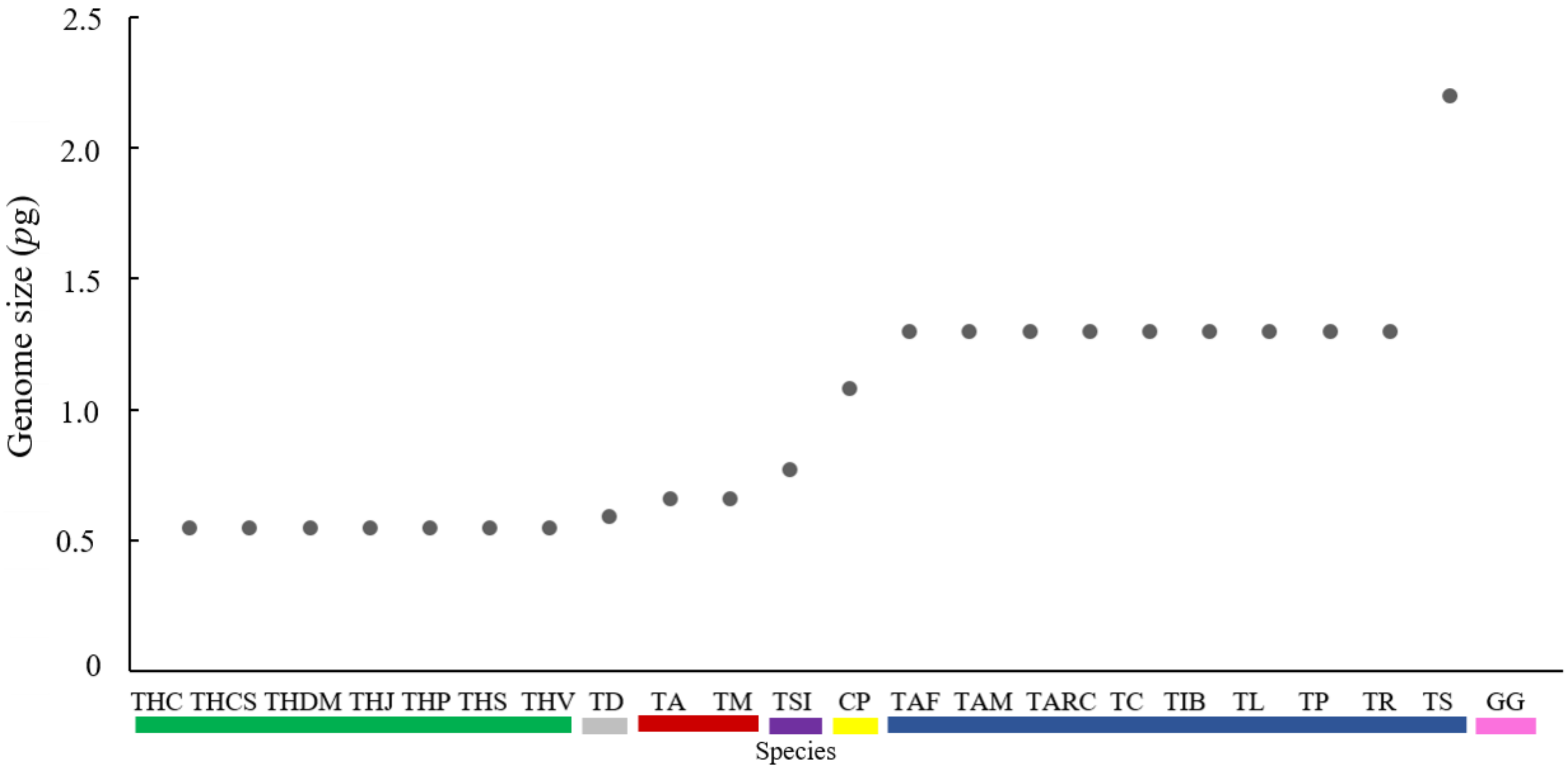


Figure S2. Mean nuclear genome size of the Cleomaceae species. The coloured bars are related to the groups observed in Figure 1. The species acronyms are presented as described in Figure 1.

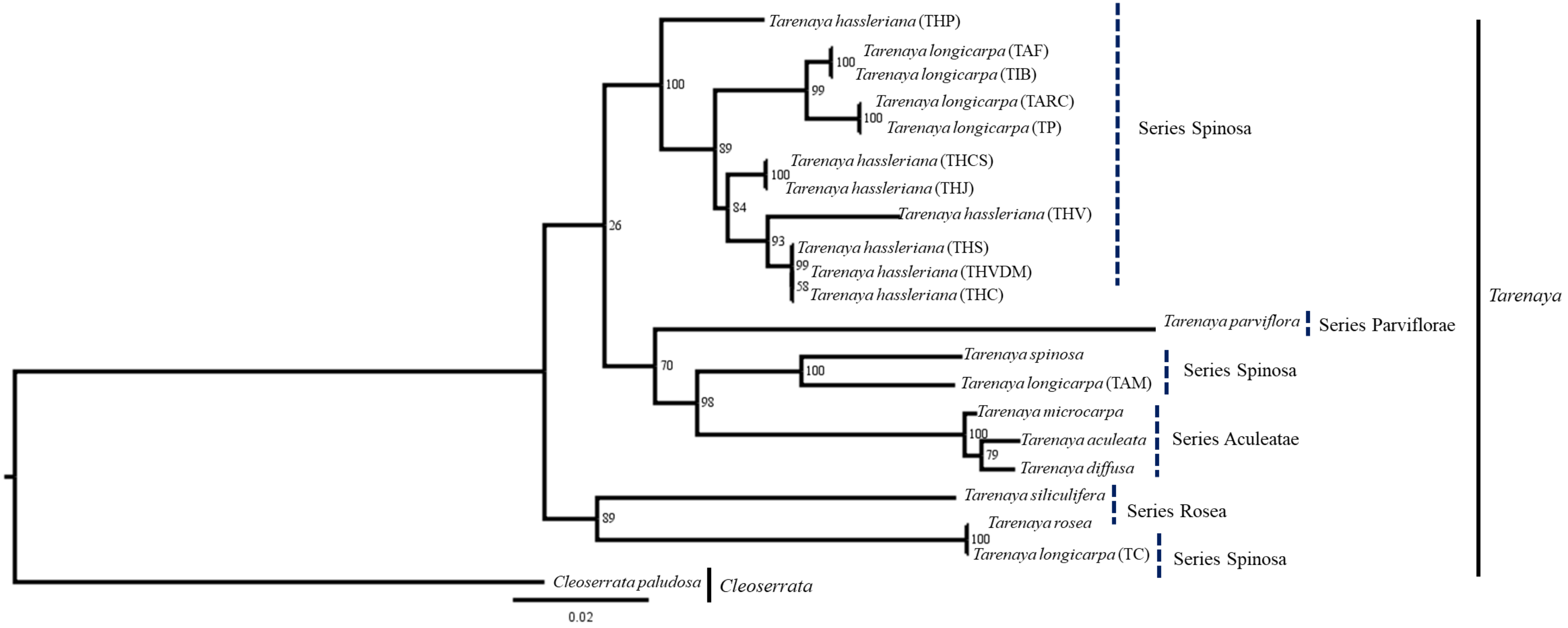


Figure S3. Molecular phylogeny of the Cleomaceae species sampled in this study. Bayesian Inference consensus tree inferred from sequences of nuclear ribossomal ITS. Numbers at nodes reflect PP. Bar: 0.02 nucleotide substitutions per site.

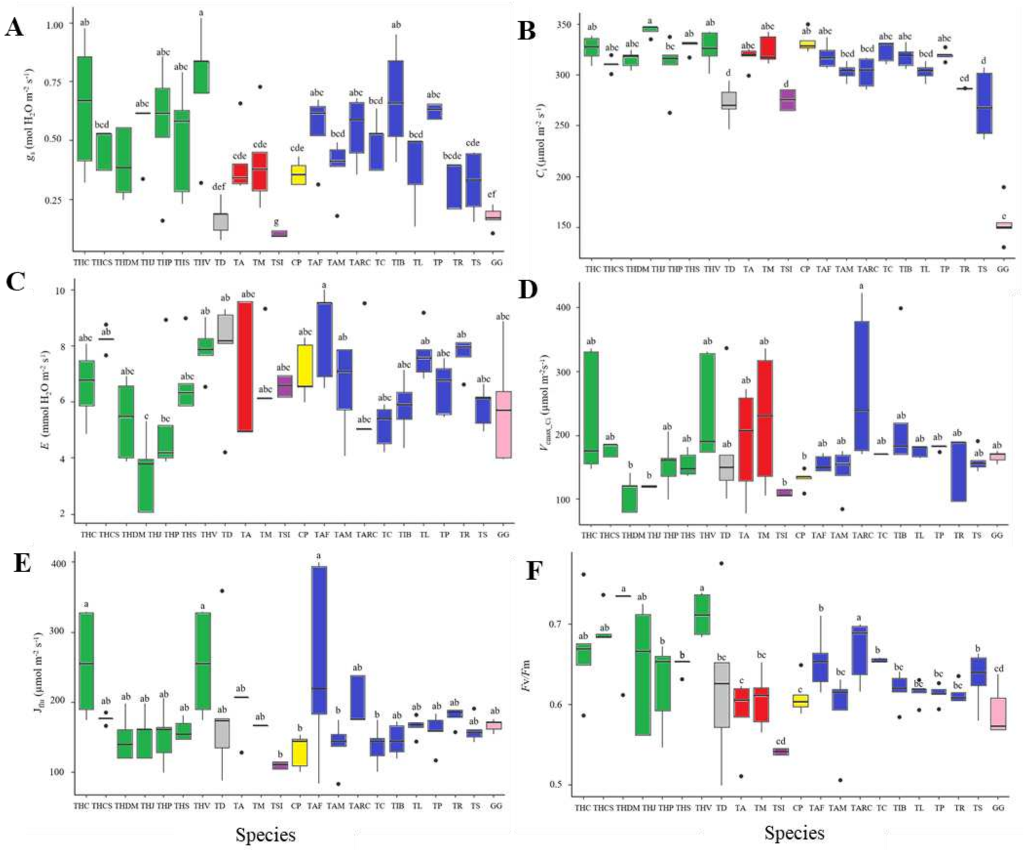


Figure S4. Gas exchange and chlorophyll *a* fluorescence parameters in Cleomaceae species. **(A)** stomatal conductance (g_s). **(B)** intern carbon (*C*i). **(C)** transpiration (*E*). **(D)** carboxylation efficiency of RuBisCO (V_{cmax_Ci}). **(E)** electron flow (J_{flu}). **(F)** PSII maximum photochemical efficiency (F_v/F_m). Letters above individual box-scatter indicate significant groupings according to Tukey's Test, n=5. The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. The coloured bars are related to the groups observed in Figure 1 as well as the species acronyms.

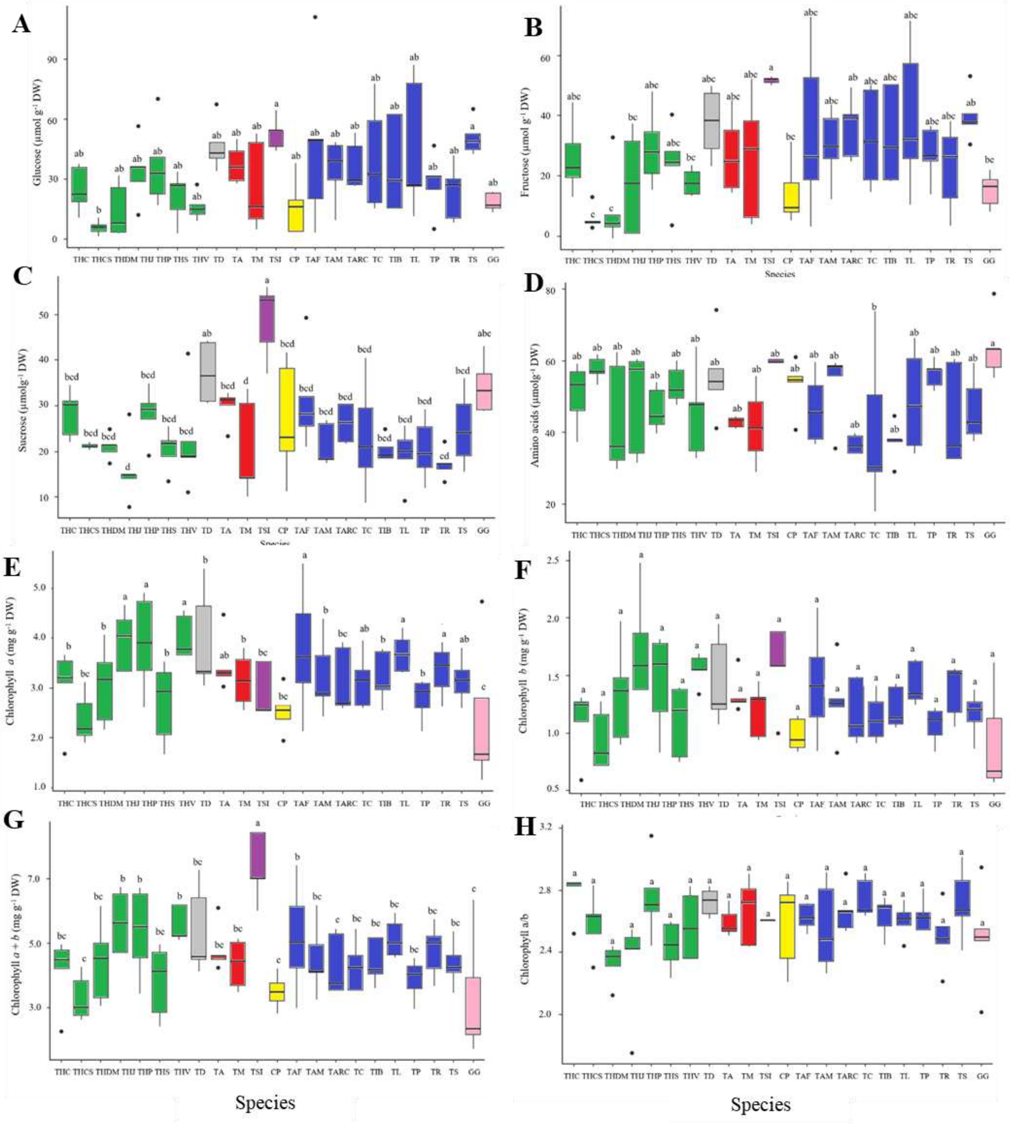


Figure S5. Levels of metabolites extracted from leaves of a subset of Cleomaceae species. For all metabolic analysis, leaf samples were harvested at middle of the day. **(A)** glucose. **(B)** fructose. **(C)** sucrose. **(D)** amino acids. **(E)** chlorophyll *a*. **(F)** chlorophyll *b*. **(G)** chlorophyll *a* + *b*. **(H)** chlorophyll *a*/*b*. Letters above individual box-scatter indicate significant groupings according to Tukey's Test (n=5). The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. The coloured bars are related to the groups observed in Figure 1 as well as the species acronyms.

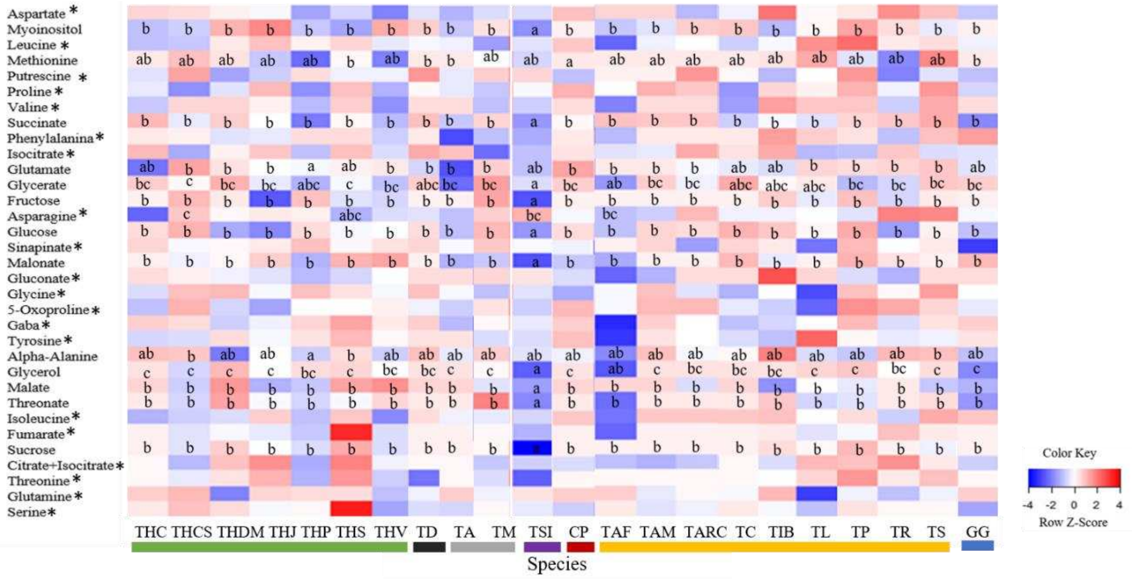


Figure S6. Heat map representing the changes in relative metabolite content in leaves collected at the middle of the day. Most species are five months old, except for TA, TM, TD and GG which are two months old. The full data sets from these metabolic profiling are additionally available in Supplemental Table 2. GC-MS data was normalized by the mean of each species/metabolite. Letters indicate significant groupings according to Tukey's Test (n=5). *Indicates not significant. The coloured bars are related to the groups observed in Figure 1A. The species acronyms are presented as described in Figure 1.

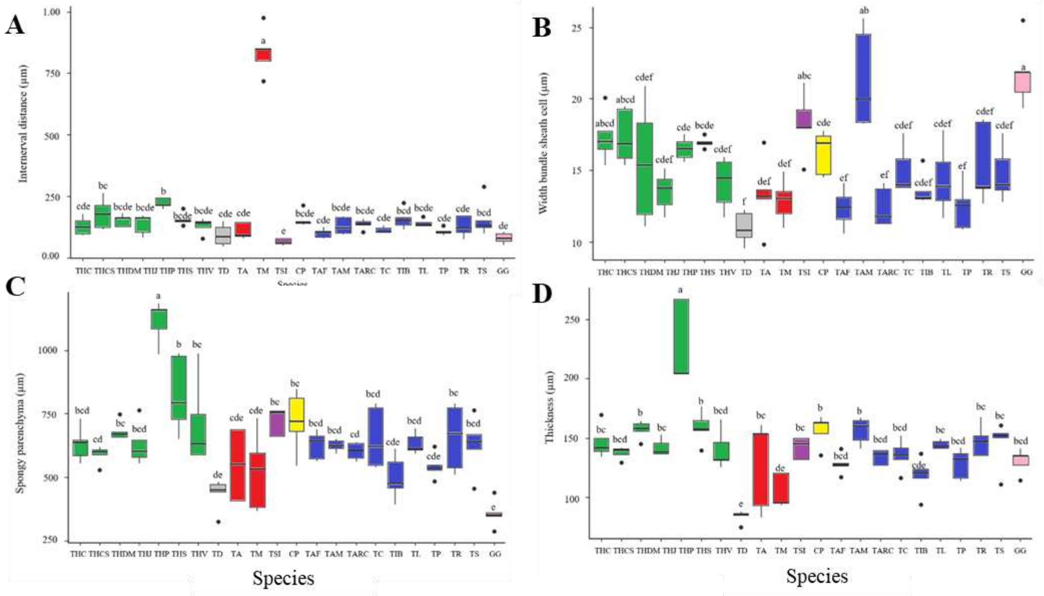


Figure S7: Natural variation of anatomical traits on leaves of a subset of Cleomaceae species. **(A)** intervenal distance. **(B)** width bundle sheath cells. **(C)** spongy parenchyma. **(D)** thickness. Letters above individual box-scatter indicate significant groupings according to Tukey's Test (n=5). The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. The coloured bars are related to the groups observed in Figure 1 as well as the species acronyms.

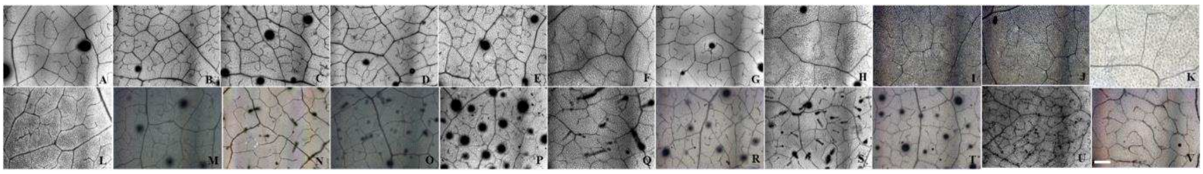


Figure S8. Vein density in leaves of Cleomaceae species. Spots indicate presence of secretory trichomes. **(A)** THC: *T. hassleriana* (Canaã-MG). **(B)** THCS: *T. hassleriana* (Canoinhas-SC). **(C)** THD: *T. hassleriana* (Domingos Martins-ES). **(D)** THJ: *T. hassleriana* (Joinville-SC). **(E)** THP: *T. hassleriana* (Piau-MG). **(F)** THS: *T. hassleriana* (São Miguel-MG). **(G)** THV: *T. hassleriana* (Viçosa-MG). **(H)** TD: *T. diffusa* (Feira de Santana-BA. **(I)** TA: *T. aculeata* (Feira de Santana-BA). **(J)** TM: *T. microcarpa* (Belém-PA). **(K)** TSI: *T. siliculifera* (Rio Pardo-MG). **(L)** CP: *C. paludosa* (Belém-PA). **(M)** TAF: *T. longicarpa* (Afrânio-PE). **(N)** TAM: *T. longicarpa* (Manaus-AM). **(O)** TARC: *T. longicarpa* (Arcoverde-PE). **(P)** TC: *T. longicarpa* (Lavras-CE). **(Q)** TIB: *T. longicarpa* (Ibimirim-PE). **(R)** TL: *T. longicarpa* (Picos-PI). **(S)** TP: *T. parviflora* (Pombal-PB). **(T)** TR: *T. rosea* (Colatina-ES). **(U)** TS: *T. spinosa* (Teresina-PI). **(V)** GG: *G. gynandra* (Mossoró-RN). The acronyms are followed by the species name and city/state of sampling state, between parenthesis. 10x. Bars = 10 µm.

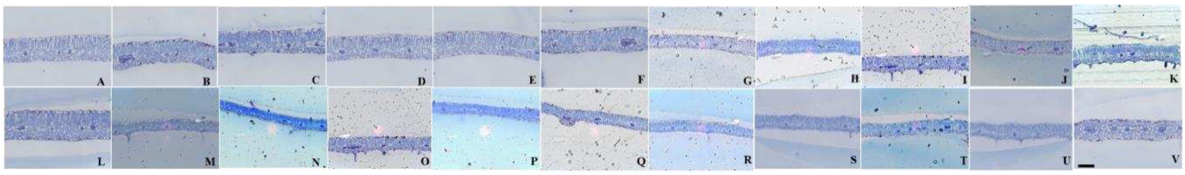


Figure S9. Leaves cross section of Cleomaceae species. **(A)** THC: *T. hassleriana* (Canaã-MG). **(B)** THCS: *T. hassleriana* (Canoinhas-SC). **(C)** THD: *T. hassleriana* (Domingos Martins-ES). **(D)** THJ: *T. hassleriana* (Joinville-SC). **(E)** THP: *T. hassleriana* (Piau-MG). **(F)** THS: *T. hassleriana* (São Miguel-MG). **(G)** THV: *T. hassleriana* (Viçosa-MG). **(H)** TD: *T. diffusa* (Feira de Santana-BA. **(I)** TA: *T. aculeata* (Feira de Santana-BA). **(J)** TM: *T. microcarpa* (Belém-PA). **(K)** TSI: *T. siliculifera* (Rio Pardo-MG). **(L)** CP: *C. paludosa* (Belém-PA). **(M)** TAF: *T. longicarpa* (Afrânio-PE). **(N)** TAM: *T. longicarpa* (Manaus-AM). **(O)** TARC: *T. longicarpa* (Arcoverde-PE). **(P)** TC: *T. longicarpa* (Lavras-CE). **(Q)** TIB: *T. longicarpa* (Ibimirim-PE). **(R)** TL: *T. longicarpa* (Picos-PI). **(S)** TP: *T. parviflora* (Pombal-PB). **(T)** TR: *T. rosea* (Colatina-ES). **(U)** TS: *T. spinosa* (Teresina-PI). **(V)** GG: *G. gynandra* (Mossoró-RN). The acronyms are followed by the species name and city/state of sampling state, between parenthesis. 10x. Bars = 10 µm.

CHAPTER 4

Harnessing genetic variation in metabolic traits to understand evolutionary pressures on Cleomaceae photosynthesis

Daniele F. Parma[1], Marcelo G. M. V. Vaz[1], David B. Medeiros[2], Auxiliadora O. Martins[1], Wagner L. Araújo[1], Andreas P. M. Weber[3] & Adriano Nunes-Nesi[1]*

[1] Departamento de Biologia Vegetal, Universidade Federal de Viçosa, 36570-900, Viçosa, Minas Gerais, Brazil.

[2] Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany.

[3] Institute for Plant Molecular and Developmental Biology, Heinrich Heine University, Düsseldorf, Germany.

*Corresponding author:
Adriano Nunes-Nesi
Departamento de Biologia Vegetal,
Universidade Federal de Viçosa,
36570-900 Viçosa, Minas Gerais, Brazil
Phone: +55-31-3612-5357
Email: nunesnesi@ufv.br

ABSTRACT

In the Cleomaceae family distinct photosynthetic mechanisms are found in different species. Thus, this group of species is of key importance to understand the evolution of carbon concentrating mechanisms. To gain more insight into the evolution of C_4 photosynthesis in this family, we examined fifteen Cleomaceae species, that have different forms of growth (e.g. herbaceous, shrub and sub-shrub), collected from different Brazilian biomes. We performed a detailed characterization of photosynthetic, anatomical, molecular and biochemical parameters in all collected species. In this regard, in addition to determination of photosynthetic parameters we have biochemical analyzes in leaves throughout the diurnal cycle. Based on these analyzes, we could separate the species into four groups: one formed by a C_4 specie (*Gynandropsis*) and three with C_3 photosynthesis (Spinosa I, Aculeatae and Spinosa II + Cleorosea + Parviflorae). Interestingly, the groups of C_3 species exhibit clear variation in anatomical, physiological and biochemical parameters, suggesting a great natural variation between these species. Furthermore, the unrecognized variation in leaf structure and physiology within the studied species indicates varying degrees of development to the C_3-C_4 photosynthetic mechanism. In addition, the studied access of *G. gynandra*, despite exhibiting physiological, biochemical and anatomical traits coherent with C_4 species, clearly differ from those found for African and Asian accessions already described, including at phylogenetic level. Considering this diversity we conclude that a group of species, especially herbaceous species, have a tendency towards a physiological, metabolic and anatomical response pattern similar to the C_4 species. Furthermore, the Brazilian accessions of Cleomaceae here characterized exhibit a high genetic diversity and might represent different stages of development from C_3 to C_4 photosynthetic mechanism.

Keywords: Brazilian biomes, *Cleoserrata*, *Gynandropsis*, photosynthetic mechanism C_4, *Tarenaya.*

INTRODUCTION

During the evolution of the C_4 carbon concentrating mechanism (CCM) from C_3 species, a combination of complex anatomical and biochemical specialization was required (Sage 2004; Gowik and Westhoff, 2011). This evolutionary process of transition from C_3 to C_4 photosynthetic mechanism occurred independently in more than 19 families and 60 angiosperms taxa, which are generally herbs and only trees in exceptionally rare cases (*e.g. Euphorbia* sp. in Hawai and *Haloxylon* sp. in Asia) (Sage 2004). Moreover, the evolution of C_4 photosynthesis is not a directional process and occurred in a series of events (Sage 2004; Gowik and Westhoff, 2011). These events include development of larger bundle sheath cells (BSC), increased leaf venation, restriction of glycine decarboxylase to BSC and establishment of photorespiratory carboxylase cycle, as well as enhancement of phosphoenolpyruvate carboxylase (PEPC) activity, establishment of C_4 cycle and optimization of the C_4 syndrome (Gowik and Westhoff, 2011; Sage, 2004).

Although all genes important for the C_4 pathway are expressed at relatively low levels in C_3 leaves (Gowik and Westhoff, 2011; Sage, 2004), the mechanism for recruitment of these genes into the C_4 pathway remains to be fully elucidated (Bergh et al., 2014). Recently, several research efforts to understand the evolution of the C_4 photosynthetic mechanism have been made (*e.g.* Bayat et al., 2018; Bergh et al., 2014; Heckmann et al., 2013; Reeves et al., 2018; Schlüter et al., 2017; Schuler et al., 2016; Williams et al., 2012). Despite the progress, important questions are still unanswered, such as the identity of transporters that ensure the increased inter- and intracellular metabolic fluxes and the genes that regulate and maintain the alterations in cell and overall leaf morphology (Gowik et al., 2011).

Noteworthy, the majority of studies on CCMs has been conducted with plants phylogenetically distant to the model *Arabidopsis thaliana* (*e.g. Flaveria* – Asteraceae) (Gowik et al., 2011; Schulze et al., 2013). In this sense, Cleomaceae, a family closely related to Brassicaceae, in which species with C_3, C_4 and intermediate C_3-C_4 photosynthetic mechanisms are found, is considered promising in studies aiming to understand the evolution of C_4 photosynthetic mechanism (Marshall et al., 2007; Voznesenskaya et al., 2007; Bayat et al., 2018; Reeves et al., 2018). In Cleomaceae, it is possible to observe three species with C_4 photosynthetic metabolism (*Gynandropsis gynandra*, *Coalisina angustifolia* and *Areocleome oxalidea*); one species with intermediate C_3-C_4 metabolism (*Coalisina paradoxa*) (names updated according to International Plant Names Index, 2020) and approximately 200 species, displaying C_3 metabolism (Feodorova et al., 2010; Marshall et al., 2007; Voznesenskaya et al., 2007; Parma et al. *submitted*).

Cleomaceae species with C_4 mechanism have different origins, which can be evidenced by both molecular phylogeny (Patchell et al., 2014) and different types of Kranz anatomy (*e.g. A. oxalidea* and *G. gynandra* have atriplicoid and *C. angustifolia*, angustifolioid kranz anatomy) (Koteyeva et al., 2011). In addition, it has also been shown that even in C_3 plants, anatomical and physiological traits indicate that some species are evolving to the C_4 photosynthetic mechanism (Marshall et al., 2007; Khoshravesh et al., 2020; Parma et al. *submitted*). Cleomaceae species apparently have remarkable flexibility in leaf development regarding C_4 metabolism (Marshall et al., 2007; Parma et al. *submitted*). In this sense, *C. violacea* possess increased venation (compared to *G. gynandra*), and others species, such *C. africana*, has more veins, larger BSC and more PEPC protein (Marshall et al., 2007), some characteristics of plants with C_4 metabolism.

Recently, it was observed that different accessions of *G. gynandra* exhibit natural variation in key characteristics of C_4 photosynthesis (Reeves et al., 2018). As consequence, the natural variation observed in *G. gynandra* can help to delineate genomic regions associated with C_4 trait variation (Reeves et al., 2018). However, it is worth to mention that it is known that C_3 plants also display variation in photosynthetic mechanisms even within the same species (see Lawson, Kramer and Raines, 2012; Rosado-Souza et al., 2015; de Oliveira Silva et al., 2018). Accordingly, this natural variation is also observed in other species of Cleomaceae, mainly in species with wide distribution and, therefore, occupying several types of environments (*e.g. Tarenaya spinosa, T. hassleriana* and *G. gynandra*) (Parma et al. *submitted*). Furthermore, it is relevant to note that this phenotypic plasticity is likely to occur due to some genetic variations (genetic diversity) (Anderson et al., 2012). Thus, the use of natural variation (genetic variability) to study the function of genes encoding for such signaling pathways may be anticipated to provide further insights into the regulation of some genes, such as C_4 photosynthesis related genes (Reeves et al., 2018). These variations may be exploited for the purpose of increasing the yield of some crops of economic interest (de Oliveira Silva et al., 2018; Hervé et al., 2001; Rosado-Souza et al., 2015; van Bezouw et al., 2019).

Accordingly, knowing the existence of a genetic variability among members of Cleomaceae and that the plants with C_4 metabolism, in the great majority, have the herbaceous habit, we aimed to investigate if the habit can influence the acquisition of C_4 type CCM related traits. For that, we perform photosynthetic response curves at different internal CO_2

concentrations and light intensities, diurnal changes in sugars levels, molecular phylogenetic analyzes involving nuclear and chloroplast markers, in addition to morphological and anatomical measurements. That way, we analyze 15 Brazilian Cleomaceae species (13 *Tarenaya* spp., 1 *Cleoserrata* sp., and 1 *Gynandropsis* sp.) that have different growth habits and are promising for understanding the evolution of the C_3-C_4 and C_4 photosynthetic mechanism (see Parma et al. *submitted*). In this regard, we highlight that a group of species, especially herbaceous species, have a tendency towards a physiological, metabolic and anatomical response pattern similar to the C_4 species. Furthermore, the Brazilian accessions of Cleomaceae here characterized exhibit a high genetic diversity and might represent different stages of development from C_3 to C_4 photosynthetic mechanism.

MATERIAL AND METHODS

Taxon Sampling

Botanical material of species belonging to Cleomaceae family was collected from different Brazilian states, which are distributed among different biomes (Fig. 1A, Table 1). In addition to the different biomes, the species were collected under different conditions, *C. paludosa*, *T. microcarpa* and most *T. hassleriana* accessions, were collected from riverbanks in abandoned fields, except the THJ accesses that was in an access road. The other species studied were collected in abandoned fields. For a subset of fifteen species of Cleomaceae selected for this study, seeds were collected *in situ*, germinated in plastic containers according to Parma et al. (*submitted*). Five plants of each species were grown in a greenhouse, with semi-controlled conditions (maximum photosynthetically active radiation of 1,500 μmol photons m^{-2} s^{-1}, with temperature of 30 ± 2 °C), being irrigated daily. Under these conditions, the plants remained for a period of two (group II – TA: *T. aculeata* (Feira de Santana -BA), TD: *T. diffusa* (Feira de Santana -BA), TM: *T. microcarpa* (Belém - PA) – and IV – GG: *G. gynandra* (Mossoró - RN) - herbaceous species) or four months (group I – TAM: *Tarenaya* sp. (Manaus - AM), TL: *T. longicarpa* (Picos - PI), TS: *T.spinosa* (Teresina - PI) – and III – CP: *C. paludosa* (Belém - PA), TP: *T. parviflora* (Pombal - PB), TR: *T. rosea* (Colatina - ES), THS: *T. hassleriana* (São Miguel - MG), THV: *T. hassleriana* (Viçosa - MG), THJ: *T. hassleriana* (Joinville - SC), THC: *T. hassleriana* (Canaã - MG), THD: *T. hassleriana* (Domingos Martins - MG) – shrubby species), comprising the spring and summer period, until physiological characterization as well as sampling of plant material for anatomical and biochemical analyzes.

DNA extraction, amplification and sequencing

Total DNA was extracted from fresh samples leaves (plants listed in Table 1) using a modified CTAB method (Doyle et al., 1990). Three genetic regions (molecular markers) were investigated: two from chloroplast – cpDNA (*ndhF*, *matK*) and, one from nucleus – nDNA (ITS), considered as phylogenetically informative at the inter- and infrageneric levels in the Cleomaceae family (Inda et al., 2008; Patchell et al., 2014; Soares-Neto et al. 2020a). The segments of interest were amplified according to Parma et al. (*submitted*). Primers used in this study were based on previous studies (Koch et al., 2001; Olmstead et al., 1993; White, 1990). The thermal cycling conditions were carried out according to Patchell et al., (2014). PCR products and generated fragments were analyzed as decribed by Parma et al. (*submitted*).

The nucleotide sequences obtained in this study (partial ITS, *matK* and *ndhF* sequences) were used to generate phylogenetic trees (Table 1). For the phylogenetic reconstruction based on concatenated gene sequences (ITS, *matK* and *ndhF*), the chosen sequences (31) were aligned and trimmed (matrix with a 4074-bp length) using *CLUSTAL W*. After, phylogenetic inferences were carried out, based on the Maximum Likelihood (ML), Bayesian inference (BI) and Maximum Parsimony (MP) methods.

The ML analysis were carried out using RAxML version 8.2.10 and node support was obtained using the rapid bootstrap algorithm with 1,000 replicates (Stamatakis, 2014). The search for the best-scoring ML tree was performed with the general time-reversible model of DNA sequence evolution with gamma-distributed rate heterogeneity (GTRGAMMA model). The MP analyzes was conducted with tree bisection-reconnection (TBR) branch swapping and the initial trees were generated by 10,000 random stepwise addition replicates. The strict consensus trees were obtained and support for each node was assessed by bootstrap analysis with 2,000 replicates through the heuristic search method. All characters were treated as unweighted and unordered. The BI analyzes was performed according to the best fit evolutionary model which was selected by the Akaike information criterion (AIC) (Posada and Buckley, 2004) using MrModeltest 2.2 (Nylander, 2004). The Markov chain Monte-Carlo algorithm was executed with four runs with 50 million generations each and sampling every 5000 generations. The first 25% of the trees were discarded as burn-in, the remaining trees were used to construct the majority rule consensus tree, and then the PP for each node was calculated.

Table 1. Description of the novel Brazilian species used in molecular phylogeny analyzes, and their respective collection sites (city-state abbreviation), herbarium vouchers and NCBI ID.

Species	**ID**	**City – State***	**Local - Environment**	**Voucher - Herbarium**	**NCBI ID**		
					ITS	***mat*K**	***ndh*F**
Cleoserrata paludosa	CP	Belém – PA	marshy region	VIC 51788	MF287989	-	MT094416
Gynandropsis gynandra	GG	Mossoró – RN	crop invasive	VIC 51789	DQ455790	KR734800	-
Tarenaya aculeata	TA	Feira de Santana – BA	open field, abandoned	VIC 51791	MT090703	-	MT094417
Tarenaya diffusa	TD	Feira de Santana - BA	open field, abandoned	VIC 51793	MT090707	MT094414	-
Tarenaya hassleriana	THV	Viçosa – MG	marshy region	VIC 51794	MT090705	MT094412	MT094420
Tarenaya hassleriana	THC	Canaã – MG	open field, abandoned	VIC 51795	MT090700	-	-
Tarenaya hassleriana	THD	Domingos Martins – ES	roadside	VIC 51907	MT090714	MT094411	-
Tarenaya hassleriana	THS	São Miguel do Anta – MG	open field, abandoned	VIC 51796	MT090709	-	MT094421
Tarenaya hassleriana	THJ	Joinville – SC	roadside	VIC 51802	MT090704	MT090713	MT094419
Tarenaya longicarpa	TL	Picos – PI	open field, abandoned	VIC 51799	MT090718	-	MT094427
Tarenaya microcarpa	TM	Belém – PA	marshy region	VIC 51797	MT090701	MT094410	MT094422
Tarenaya parviflora	TP	Pombal – PB	marshy region	VIC 51798	MT090710	-	-
Tarenaya rosea	TR	Colatina – ES	open field, abandoned	VIC 51800	MT090702	-	MT094423
Tarenaya sp.	TAM	Manaus – AM	open field	VIC 51803	MT090707	-	-
Tarenaya spinosa	TS	Teresina - PI	roadside	VIC 51804	MT090706	MT094409	MT094425

*States: PA – Pará; RN – Rio Grande do Norte; BA – Bahia; MG – Minas Gerais; ES – Espírito Santo; SC – Santa Catarina; PI – Piauí; PB – Paraíba; AM – Amazonas.

Micromorphology analysis

Anatomical survey was conducted to evaluate the venation density of fully developed leaves of the upper third. Diaphanization of the plant material was carried out according to the methodology previously described (Zsögön et al. 2015). Subsequently, samples were assembled on glass slides and images of the adaxial epidermis were obtained with the aid of a Zeiss Axio Scope A1 photomicroscope coupled to a color image capture system (Axiovision® 105). The rib density was defined as the measurement of the lengths of 2nd, 3rd and 4th order veins in a given leaf area (mm mm^{-2}). Quantification was performed using Image Pro-Plus® software (version 4.5, Media Cybernetics, Silver Spring, USA).

In addition, fully expanded leaves from the third node were collected and fixed in formaldehyde: acetic acid: ethanol 50% (FAA50) at a ratio of 5:5:9 (v/v) for 48 hours. After, the material was dehydrated in ethanolic series and included in historesin (Leica, Heidelberg, Germany). The material was cut transversely (5 µm) in a self-advancing rotary microtome (RM 2155, LEICA, Heidelberg, Germany) using glass razors. The sections were stained with 0.05% toluidine blue in acetate buffer (pH 4.7) (O'Brien et al., 1964) for one minute, and assembled between slides and coverslips with synthetic resin. The photomicrographs were obtained in a photomicroscope (Zeiss, MC-80), with the micrometric scales photographed and enlarged under the same optical conditions. Quantification and measurements of leaf thickness, palisade parenchyma, spongy parenchyma, intervein distance and width bundle sheath cell were performed using Image Pro-Plus® software.

Gas-exchange and fluorescence analyzes

Fully expanded leaves from the third node in non-flowering plants, two-month herbaceous and four-month-old shrub, were used for gas-exchange measurements with a LICOR 6400XT (LI-COR Biosciences, Lincoln, USA). The measurements were carried out with a flow rate of 400 µmol CO_2 s^{-1}, light intensity of 1,000 µmol photons m^{-2} s^{-1}. All measurements were performed using the 2 cm^2 leaf chamber at 25°C, and the leaf-to-air vapor pressure deficit was kept at 1.2 to 2.0 kPa, while the amount of blue light was set to 10% of the PPFD to optimize stomatal aperture.

An initial analyzes of the Cleomaceae species and comparison with the associated C_3 plants were carried out through photosynthesis response curves (A_n) at different internal CO_2 concentrations (C_i). The curves were initiated at internal CO_2 concentration of 400 µmol photons mol^{-1} and photosynthetic photon flux density (PPFD) of 1,000 µmol photons m^{-2} s^{-1}. After, the *Ci* was reduced to 0 µmol CO_2 mol^{-1} and then gradually increased to 1,600 µmol CO_2 mol^{-1}; in total 15 different CO_2 concentrations were tested. Simultaneously, chlorophyll *a* fluorescence parameters were quantified (Yin et al., 2009). Corrections for the leakage of CO_2 into and water vapor out of the leaf chamber of the LI-6400 were applied to all gas-exchange data as previously described (Rodeghiero et al., 2007).

The initial fluorescence (F_0) was measured by illuminating dark-adapted leaves (1 h) with weak modulated measuring beams (0.03 µmol m^{-2}·s^{-1}). A saturating white light pulse (8000 µmol photons m^{-2}·s^{-1}) was applied for 0.8 s to obtain the maximum fluorescence (F_m), from which the variable-to-maximum chlorophyll fluorescence ratio was then calculated: $F_v/F_m = [(F_m- F_0)/F_m)]$. In light-adapted leaves, the steady-state fluorescence yield (F_s) was measured with the application of a saturating white light pulse (8000 µmol m^{-2}·s^{-1}) to achieve the light adapted maximum fluorescence (F_m'). A far-red illumination (2 µmol m^{-2}·s^{-1}) was applied after the actinic light be turned off to measure the light-adapted initial fluorescence (F_0').

The capture efficiency of excitation energy by open PSII reaction centers (F_v'/F_m') was estimated following Logan et al., (2007) and the actual PSII photochemical efficiency (Φ_{PSII}) was estimated as (Φ_{PSII}) = ($F_m'– F_s'$)/F_m' (Genty et al., 1989). As the Φ_{PSII} represents the number of electrons transferred per photon absorbed in the PSII, the electron transport rate (J_{flu}) was calculated as $J_{flu} = \Phi_{PSII}\cdot\alpha\cdot\beta\cdot PPFD$, where α is leaf absorptance and β reflects the partitioning of absorbed quanta between PSII and PSI, and the product αβ was adopted as described in the literature for *Arabidopsis* as equal to 0.451 (Flexas et al., 2007).

The photosynthetic light response curves (A_N/PPFD) were initiated at a concentration of CO_2 (C_a) of 400 µmol CO_2 mol^{-1} and PPFD of 1,000 µmol photons m^{-2} s^{-1}. Then, the PPFD were increased to 1,200 µmol m^{-2} s^{-1} up to 1,600 µmol m^{-2} s^{-1} and then gradually decreased to 0 µmol m^{-2} s^{-1}; a total of 13 different PPFDs were used. Dark respiration (R_d) was measured using the same gas-exchange system as described above, after at least 1 h in the dark period. The obtained values were divided by two (R_d/2) to estimate the mitochondrial respiration rate in the light.

The concentration of CO_2 in the carboxylation sites (C_c) was calculated following the equation described by Harley et al., (1992):

$$C_c = (\Gamma^*(J_{flu} + 8(A_N + R_L))/(J_{flu} - 4(A_N + R_L)$$

where the conservative value of Γ^* for *Arabidopsis* was taken from (Walker and Cousins, 2013). Then, mesophyll conductance (g_m) was estimated as the slope of the A_N versus $C_i - C_c$ relationship ($C_i < 300$ µmol·mol^{-1}).

Given that current methods for estimating g_m include several assumptions as well as technical limitations and sources of error that need to be considered to obtain reliable values (Pons et al., 2009), we estimated this parameter by the Ethier and Livingston, (2004) method, which fits A_N/C_i curves with a nonrectangular hyperbola version of Farquhar-von Caemmerer-Berry FvCB model, based on the hypothesis that g_m reduces the curvature of the RuBisCO-limited portion of an A_N/C_i curve. From A_N/C_i and A_N/C_c curves, the maximum carboxylation velocity (V_{cmax}) and the maximum capacity for electron transport rate (J_{max}) were calculated by fitting the mechanistic model of CO_2 assimilation (Farquhar et al. 1980) using the C_i or Cc based temperature dependence of kinetic parameters of RuBisCO (K_c and K_o; Walker and Cousins, 2013). Then V_{cmax}, J_{max}, and g_m were normalized to 25°C using the temperature response equations from (Sharkey et al., 2007).

The photosynthetic limitations were calculated based on the approach described by Grassi and Magnani, (2005). This method uses the values of A_N, g_s, g_m, V_{cmax}, Γ^*, C_c, and $K_m = K_c\ (1 + O/K_o)$ and permits the partitioning into the functional components of photosynthetic constraints related to stomatal (l_s), mesophyll (l_m), and biochemical (l_b) limitations.

Metabolite Analyzes

Fully expanded leaves of the third node were harvested at beginning, middle and end of light period in non-flowering plants, soon after the photosynthetic parameters analyzes. At this time, herbaceous plants were two months old and the other shrub species were four months old. Metabolite extraction was performed by grinding the material with liquid nitrogen and addition of the appropriate extraction buffers. Approximately 25 mg of leaf freeze-dried matter were subjected to ethanolic extraction as previously described (Cross et al., 2006).

The concentrations of glucose, fructose and sucrose were determined from the supernatant as previously described (Fernie et al. 2001), as well as total amino acids (Cross et al. 2006), malate and fumarate (Nunes-Nesi et al. 2007). From the ethanol insoluble fraction, the concentrations of starch (Fernie et al. 2001) and total soluble proteins contents (Bradford 1976) were determined by established protocols.

Statistics Analyzes

Data were obtained from five plants per species, each of them was placed in individual pots, and each pot representing a biological replicate. The experiment was designed in a completely randomized distribution. The effect of species was determined by analyzes of variance ($P < 0.05$). The means were submitted to the Tukey test. To evaluate the relations between the different species, we also used factor analyzes and, subsequently, principal component analyzes (PCA) based on the correlation matrix for the adjusted means (score Z). All statistical analyzes were carried out using Statistica Software and R.

RESULTS

Phylogenetic analyzes of Brazilian Cleomaceae species

Considering the molecular phylogeny performed in this study, based on the ITS, *mat*K and *ndh*F partial sequences, the 15 Cleomaceae morphospecies were grouped into three genera (*Cleoserrata*, *Gynandropsis* and *Tarenaya*) (Fig. 1B), that are presented as monophyletic groups. In addition, the *Tarenaya* series/sections, the genus with the largest number of species analyzed here (13 spp.), also presented as monophyletic (Spinosa, Aculeatae, Parviflorae and Cleorosea).

The phylogenetic position of the studied accessions of Cleomaceae is also corroborated by morphological evaluation (Fig. 1C), but it is not congruent with the grouping performed by the analysis of main components (Fig. 1D) demonstrating the physiological and metabolic diversity of the Brazilian Cleomaceae species. On this subject, *Gynandropsis* and Aculeatae groups are formed by herbaceous species and the clusters Spinosa I, Spinosa II, Parviflorae and Cleorosea are composed by species with shrub habit, while Cleoserrata is subshrub (Fig. 1C). The groups were presented as monophyletic for all concatenated markers and all three methods used (Maximum parsimony, Maximum likelihood and Bayesian) (Fig. 1B), which suggests that for Cleomaceae the three markers have a good phylogenetic signal and can be used individually. Although the topology of the phylogenetic tree is similar among the different methods, the Bayesian inference showed a better phylogenetic signal for the group, being, therefore, the most recommended to evaluate the phylogenetic proximity between the species of Cleomaceae (Fig. 1B).

Principal component analysis of Brazilian Cleomaceae species based on physiological, anatomical and biochemical traits

In order to understand the grouping of metabolic and physiological traits of the Brazilian Cleomaceae species described below and verify the congruence between this pattern and those found through molecular phylogeny, we carried out principal component analysis (PCA) using the first two components, which cover most of the data set variance (71.17% - Factor 1: 13.37% and Factor 2: 27.80%) (Fig. 1D). We first performed a factor analysis with all the photosynthetic, metabolic, anatomical and morphological variables determined for all 15 accessions. Accordingly, 17 variables were selected as more significant: the levels of amino acids, malate, fructose and chlorophyll *a*/*b* ratio in leaves at middle of the light period, photosynthetic CO_2 assimilation rate (A_N), CO_2 compensation point (CO_2, CP), flux carbon (A_{GROSS}), stomatal conductance (g_s), intrinsic water use efficiency (WUEi), intern carbon concentration (C_i), maximum carboxylation capacity based in C_i (V_{cmax_Ci}), maximum capacity for electron transport rate based in C_i (J_{max_Ci}), mesophyll (ML) and biochemical limitation (BL) as well as width bundle sheath cells (WBSC) and spongy parenchyma (SP), leading to four major plant-grouping (Fig. 1D).

Interestingly, the group I is composed of Cleomaceae species with larger height (Fig. 1C). Moreover, in relation to plant architecture, the species of group I are morphologically similar to each other, distinguished mainly by the amount and location of spines and trichomes and largest fruit production (Supplementary Fig. 2A-D). Regarding the group II that comprises species of herbaceous habit, besides producing smaller fruits and lower number of seeds per fruit, they produce largest seeds with aryl compared to the species from the other groups. The group III, in turn, is the most morphologically diverse, which can be observed by leaf morphology, mainly, number of leaflets (3-7), presence/absence of spines and hispid trichomes, medium fruits and intermediate height (Supplementary Fig. 2A-D). It is noteworthy that species from groups I, II and III display C_3 photosynthetic mechanism, while *G. gynandra*, that clustered solely in group IV has C_4 photosynthetic mechanism. In addition, this accession is an herbaceous plant with *ca.* 1 m height that produces the smallest seeds and fruits having velutine trichomes.

Photosynthetic parameters of Brazilian Cleomaceae species

In order to fully characterize and understand the physiological basis for the differences between the four groups of Cleomaceae (Fig. 1D) we performed detailed physiological analyses (Fig. 2A-H) by diffusional photochemical and biochemical constraints to photosynthesis. Analyzing gas-exchange under photosynthetically active photon flux density (A_{PPFD}) that ranged from 0 to 1,500 µmol photons $m^{-2} \cdot s^{-1}$, we observed that the groups of species exhibit different photosynthetic performance. Group IV displayed a better performance on the curve, since A_N gradually increased with higher irradiance. A similar pattern was observed in plants from the groups I and III, although with A_N slightly lower than group IV. Finally, group II tends to A_N stabilize in 500 µmol photons $m^{-2}s^{-1}$, different from the others (Fig. 2A).

Under ambient conditions, it was observed a natural variation among the Cleomaceae species concerning the physiological parameters related to photosynthesis. The species belonging to the group I, which are shrubbery, exhibited higher photosynthetic rates (A_N) coupled to higher stomatal conductance (g_s), CO_2 compensation point, as well as transpiration and respiration rates (Fig. 2A-D). In contrast, the group II, composed by herbaceous species, these parameters were significantly lower (Fig. 2A-D). Noteworthy, most of the evaluated physiological variables (A_N, gs, E, V_{cmax} and WUEi) plant from the group composed of herbaceous species (groups II and IV) exhibited a similar response pattern, as well as the shrub/subshrub species (groups III and I) have the same response pattern (Fig. 2C-I).

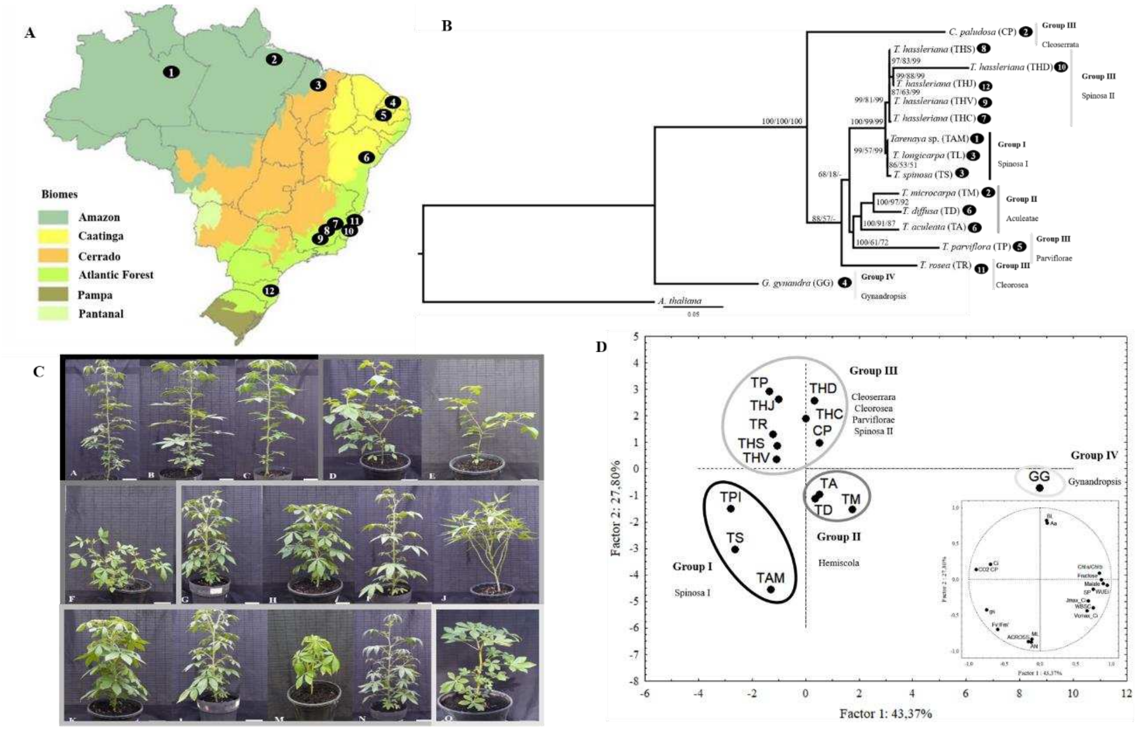


Figure 1. Map, molecular phylogeny, phenotype and PCA of Cleomaceae species. A- Map showing the sites from which 15 Cleomaceae species were collected. B- Phylogenetic relationships between Cleomaceae species by Bayesian inference of concatenated sequences of nuclear ITS and chloroplastic *matK* and *ndhF*. Numbers at nodes reflect Posterior Probability (PP) from Bayesian inference, Bootstrap values from Maximum Likelihood (ML) and Bootstrap values from Maximum Parsimony (MP): PP/ML/MP. Bar: 0.05 substitutions per nucleotide. C- Morphospecies studied [Species name (City of Sampling-state)]: A- TAM: *Tarenaya* sp. (Manaus-AM); B- TS: *T. spinosa* (Teresina-PI); C- TL: *T. longicarpa* (Picos-PI); D- TM: *T. microcarpa* (Belém-PA); E- TD: *T. diffusa* (Feira de Santana-BA); F- TA: *T. aculeata* (Feira de Santana-BA); G- THV: *T. hassleriana* (Viçosa-MG); H- THS: *T. hassleriana* (São Miguel-MG); I- TR: *T. rosea* (Colatina-ES); J- CP: *C. paludosa* (Belém-PA); K- THC: *T. hassleriana* (Canaã-MG); L- THD: *T. hassleriana* (Domingos Martins-ES); M- THJ: *T. hassleriana* (Joinville-SC); N- TP: *T. parviflora* (Pombal-PB); O- GG: *G. gynandra* (Mossoró-RN). Species name (City of Sampling-state). D- Principal component analysis and factor analysis based on 17 morphological, anatomical and metabolic data.

The Brazilian accession of *G. gynandra* (group IV) exhibited a reduced CO_2 compensation point (~ 4 ppm CO_2) as well as low transpiration rates (Fig. 2C-D) in comparison with the other groups. In addition, high RuBisCO carboxylation rate (V_{cmax}) and intrinsic water use efficiency (WUEi) were observed for the same accession (Fig. 2F-G). Interestingly, some species from groups I and II exhibited lower CO_2 compensation point (~ 50 ppm CO_2) and V_{cmax} values close to those found for the accession of *G. gynandra* (Fig. 2C and F). In addition, the different accesses of *T. hassleriana* (group III) showed different responses to the analyzed physiological variables (Supplementary Table 1).

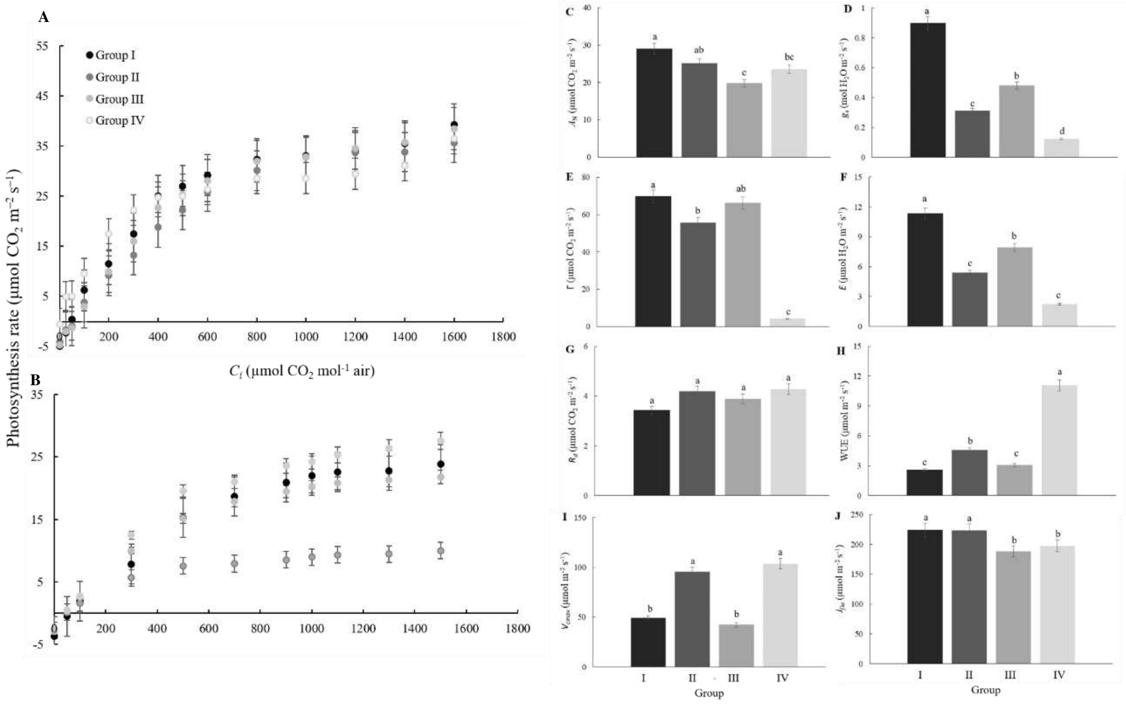


Figure 2. Net photosynthesis (A_N) curves in response to sub stomatal (C_i) CO_2 concentrations and A_{PPFD} as well as chlorophyll *a* fluorescence parameters in brazilian Cleomaceae species. Leaves from the third node, from plants at vegetative stage aged two months (groups II and IV) and four months (groups I and III) were used. A- A_N/C_i. B- A_N/A_{PPFD}. C- Ambient CO_2 assimilation rates (A_N) (400 ppm atmospheric [CO_2]). D- stomatal conductance (g_s). E- CO_2 compensation point (Γ). F- transpiration (E). G- dark respiration (R_d). H- water use efficiency (WUE). I- carboxylation efficiency of rubisc (V_{cmax_Ci}). J- electron flow (J_{flu}). Letters above individual box-scatter indicate significant groupings according to Tukey's Test, n=5. Groups I, II, III and IV were designated according to PCA – (Fig. 1).

Leaf metabolite levels of Brazilian Cleomaceae species

In order to understand the natural variance observed between the different species of Brazilian Cleomaceae we examined the level of the main carbon and nitrogen containing metabolites along the light period (Supplementary Fig. 1A-H). During the light period, the species accumulated more glucose, fructose and starch, as well as total amino acids and protein contents. Sucrose peaked in the middle of light period, which also occurred for glucose (groups III and IV). In turn, the levels of chlorophyll *a* and *b* decreased throughout the day.

When analyzing the metabolites in the middle of the light period, we observed a pattern between groups (herbaceous *versus* shrub) similar to that observed for the physiological variables, for almost all the metabolic variables presented (except chlorophyll *a*, malate and fumarate). In this sense, we observed that for glucose (Fig. 3A), groups I and III, had higher amounts, in turn in fructose (Fig. 3B), groups III and IV showed higher content, and regarding the levels of sucrose (Fig 3C) plants from the group III exhibited the largest quantity and group IV the smallest. In general, plants from groups I and III exhibited higher amounts of total sugar contents than groups II and IV. In terms of starch metabolism, we observed that starch levels (Fig. 3D) and starch degradation rate (Fig. 3E), the group IV showed a higher content of this metabolite and degradation rate, followed by group II, III and I. For nitrogenous compounds, group III had greater levels of total free amino acids (Fig. 3F) and group IV less. This behavior is opposite to that observed for total proteins levels (Fig. 3G) in which group IV had a greater amount and group III lower. As for chlorophyll *a* (Fig. 3H), we did not observe any significant difference between groups, but the chlorophyll *b* content (Fig. 3I) was higher for group I, followed by groups II, III, and IV. For malate, an important organic acid for C_4 and C_3 metabolism, it was observed that group IV had the highest content, about two times higher than the observed for the other species (Fig. 3J). Curiously, fumarate did not exhibit any clear patter between the groups of Cleomaceae (Fig. 3J).

Leaf anatomical parameters of Brazilian Cleomaceae species

Anatomical parameters are important criteria for classifying species according to the different photosynthetic mechanisms, mainly considering the Kranz anatomy observed in C_4-species. The accession of *G. gynandra,* different from the other groups, has a typical Kranz anatomy, characterized by enlarged BSC and mesophyll cells forming a ring around each BSC (Supplementary Fig. 4O). Thus, unlike the other physiological and metabolic variables described above, by anatomical parameters, we were unable to separate groups I, II and III. Thus, we have group IV responses that differ from the others, in most cases analyzed.

We evaluated the venation density in a fully expanded leaf of the third node, which has been also used as criterion to evaluate different photosynthetic mechanisms (Sage, 2004). Based on this trait, we observed that species from groups I, II and III have a higher venation density compared to *G. gynandra* (Fig. 4A; Supplementary Fig. 3A-O). In turn, *G. gynandra* has shorter intervein distance (Fig. 4C), smaller thickness (Fig. 4D) and palisade parenchyma (Fig. 4E) and greater thickness of the spongy parenchyma (Fig. 4F; Supplementary Fig. 4A-O) with values similar to those observed by Marshall et al., (2007) and Voznesenskaya et al., (2007), but different of Reeves et al., (2018). Concerning the leaf thickness, *C. paludosa* exhibited higher leaf thickness compared to other species and no pattern was observed between the groups (Fig. 4D). Noteworthy, *C. paludosa* is the only species in our population with leathery leaf; being the leaves of the others studied species classified as membranaceous (Fig. 1C).

DISCUSSION

Genotypic diversity and phenotypic plasticity of Cleomaceae species

Genetic diversity in the same species has been reported for Cleomaceae species (see Reeves et al., 2018; Sogbohossou et al., 2019; Parma et al., *submitted*). In this sense, as observed to *G. gynandra* (Reeves et al., 2018; Sogbohossou et al., 2019), we also found that accessions of *T. hassleriana* have great genetic diversity in terms of phylogenetic placement, as well as physiological, biochemical and anatomical parameters (Supplementary Table 1). Accessions of *T. hassleriana* have differences in fully expanded leaflets, which vary in size and shape, and there were variations in petiole length, presence of trichomes, and flower anthocyanin pigmentation (off-white to pink). This intraspecific variation is related the selective pressures on traits within populations (Flood, 2019) which often comes from diverse environments and/or ecological extremes (*e.g*. Yeaman et al. 2016; Hancock et al. 2011).

In fact, the risk of colonizing exposed habits depends on the actual rate of *E,* being determined by the coupling of micrometeorological effects with the g_s (Osborne and Sack, 2012). In this respect, the g_s is controlled by changes in stomatal aperture, which are regulated by leaf water status, A_N and chemical signals reflecting the soil water status transmitted from roots to leaves (Buckley 2005). In this way, the water relations have previously been proposed by Osborne and Sack (2012) to influence the evolution of C_4 photosynthesis in an indirect way. Thus, the g_s plays a pivotal role in leaf gas exchange, limiting both the efflux of water and influx of CO_2 (Osborne and Sack 2012). In this sense, reduced stomatal aperture restricts the CO_2 supply to photosynthesis and decreases transpiration (observed for groups II and IV), thereby reducing latent heat loss and raising leaf temperature. However, the C_4 CCM causes a large

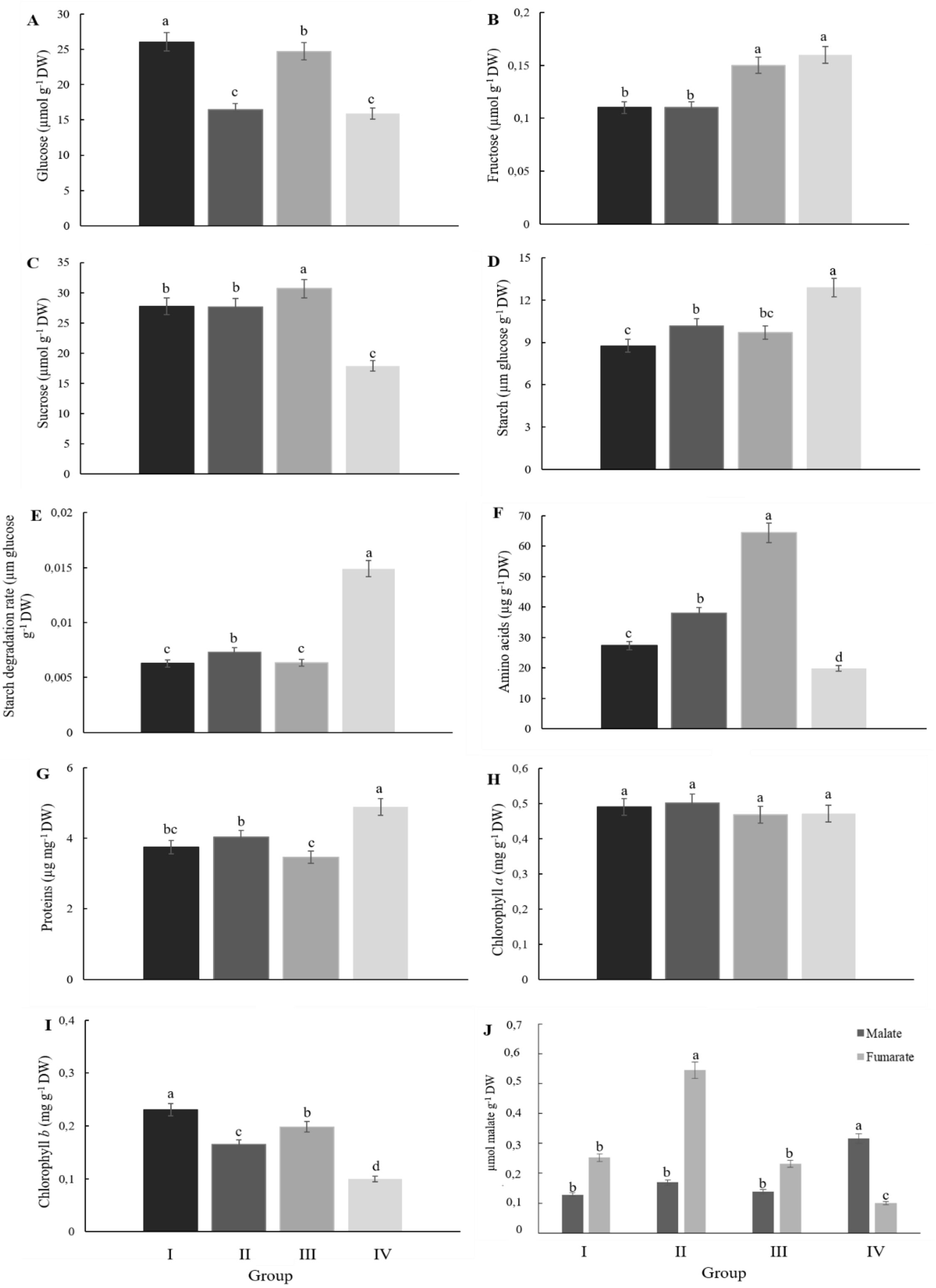


Figure 3. Levels of metabolites extracted from leaves of a subset of Cleomaceae species. For all metabolic analysis, leaf samples were harvested at the middle of the day. Leaves from the third node, from plants in vegetative stage aged two months (groups II and IV) and four months (groups I and III) were used. A- glucose. B- fructose. C- sucrose. D- starch. E- starch degradation. F- proteins. G- amino acids. H- chlorophyll *a*. I- chlorophyll *b*. J- malate and fumarate. Letters above individual bars indicate significant groupings according to Tukey's Test, n=5.

shift in this trade-off between carbon gain and water loss (Osborne and Sack 2012). In this sense, even though *G. gynandra* (group IV) has a lower value of g_s, the net rate of leaf photosynthetic CO_2 uptake is greater than C_3 species (group II).

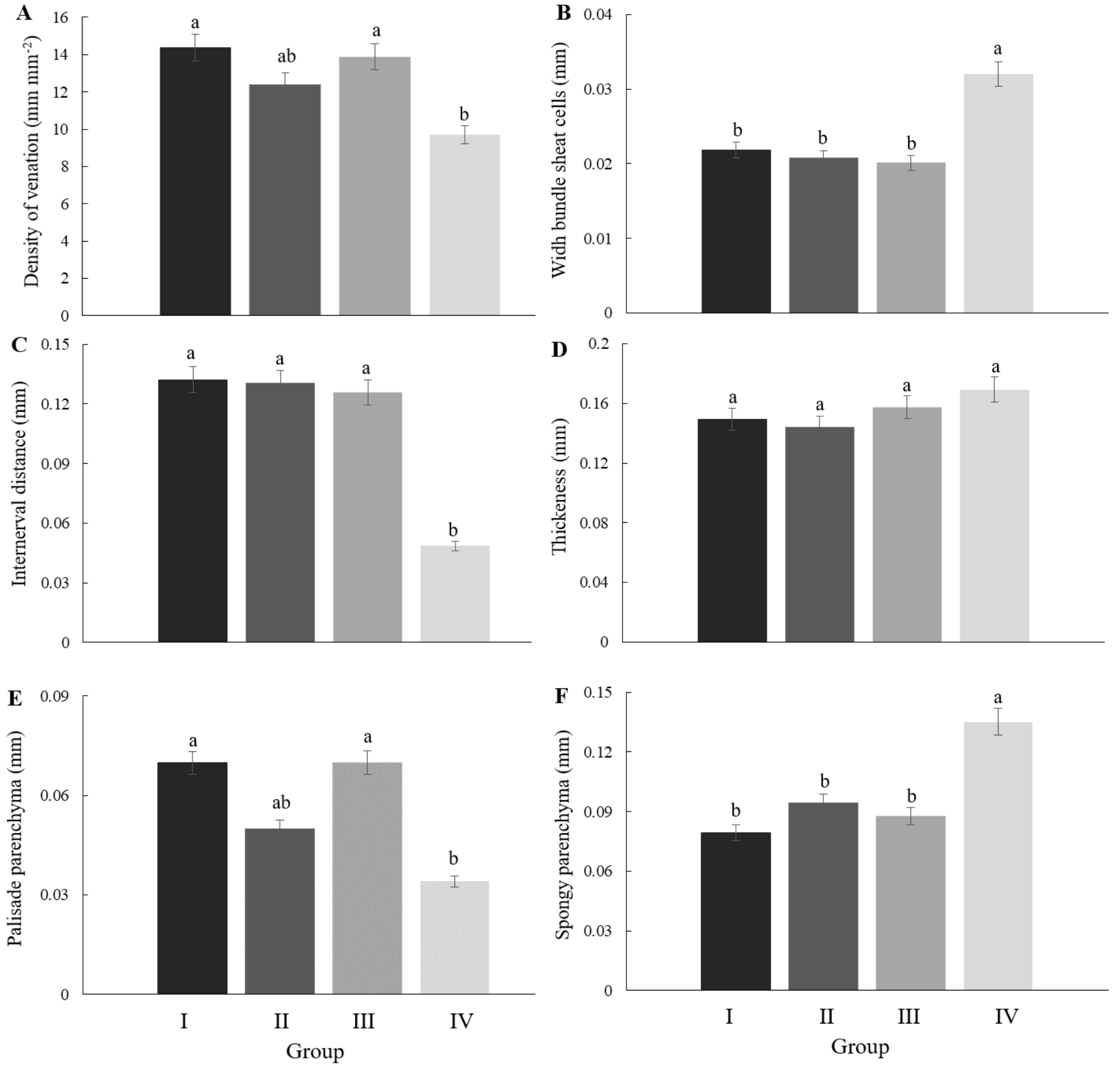


Figure 4. Natural variation of anatomical traits of leaves in a subset of Cleomaceae species. Leaves from the third node, from plants in vegetative stage aged two months (groups II and IV) and four months (groups I and III) were used. A- density of venation. B- width bundle sheath cells. C- intervenal distance. D- thickness. E- palisade parenchyma. F- spongy parenchyma. Letters above individual bars indicate significant groupings according to Tukey's Test, n=5.

Are Brazilian accessions of Cleomaceae in process of evolution from C_3 to C_4 photosynthetic metabolism?

The evolutionary path of photosynthesis C_3 to C_4, via intermediates C_3-C_4, is framed in seven distinct phases (see review by Sage 2004). These events include since increased leaf venation and development of larger BSC until enhancement of phosphoenolpyruvate carboxylase activity and establishment of C_4 cycle (Sage 2004; Gowik and Westhoff, 2011). In this way, combined morphological changes, physiological and molecular as well as environmental conditions play an important role in the development of C_4 photosynthetic mechanism. Therefore, the anatomical preconditioning required for C_4 is precisely that which is expected to evolve in C_3 plants under selection for greater tolerance to dry soil, open environments, high transpiration and low CO_2 concentration (Osborne and Sack, 2012; Zhou et al., 2018). In this sense, a direct role in water relations, as well the effects of atmospheric CO_2 depletion, provides a clear explanation for many of the anatomical changes in the early evolution of intermediate C_3-C_4 (Osborne and Sack, 2012). Therefore, the higher vein density, for example, allow greater g_s and A_N during periods with high water availability, which compensate the low atmospheric CO_2 (Osborne and Sack, 2012).

Interestingly, *Tarenaya* sp. (group I – C_3) and *T. rosea* (group III – C_3) exhibit high venation density and large BSC, besides having, together with *T. longicarpa* (group I – C_3) and *T. microcarpa* (group II – C_3), shorter intervenal distance (after *G. gynandra* – group IV – C_4). It is worth mentioning that a reduction in the distance between the leaf veins and an increase in the BSC can evolve under dry environmental conditions to improve the status of water in the leaf (Mckown and Dengler, 2007; Sage, 2004, 2001). Nonetheless, this characteristics can be called of proto-Kranz anatomy and occurs in C_3 species that are closely related to intermediate C_3-C_4 (Muhaidat et al., 2011).

Furthermore, the Brazilian accessions of Cleomaceae exhibited a CO_2 compensation point similar to those observed for species recognized as having the C_3 photosynthetic mechanism (Schlüter et al., 2017), except *G. gynandra*, that belongs to group IV and clearly exhibited photosynthetic mechanism C_4. Despite that, the group II exhibited CO_2 compensation point of ~50 ppm CO_2, that is considered low for C_3 Cleomaceae (Marshall et al., 2007; Voznesenskaya et al., 2007) and higher WUEi compared to other C_3 species (group I and III). In this sense, it is known that the evolution of a larger WUEi is apparently an important step in the evolution of C_4 photosynthesis (Osborne and Sack, 2012). Furthermore, this higher WUEi observed for group II (herbaceous C_3), which loses in values only for species C_4, may have occurred due to the low g_s (as well as for species C_4). Some groups of plants with intermediate photosynthetic mechanism C_3-C_4, improve WUE by operating a photorespiratory pump by increasing A_N to a given g_s (*e.g. Heliotropium* - Vogan et al. 2007).

It has been suggested that the size of the genome can influence/favor the acquisition of the C_4 mechanism (Parma et al. *submitted*; Bianconi et al. 2018; Christin et al., 2013; Monson, 2003; Roddy et al., 2020; Wang et al., 2009). In this way, as previously demonstrated, the genome size acts as a first-order constraint on carbon gain and is predicted to define the upper limits of allocation to growth, reproduction, and defense (Roddy et al., 2020). However, although the size of the genome is important, the gene copy number decreases as beneficial mutations in the promoter or coding sequences are fixed, in a process of neofunctionalization (Bianconi et al., 2018). Thus the beneficial effects of gene duplication for physiological innovation are therefore likely to be transitory, with no footprint on longer evolutionary scales.

In this way, the strong effects of genome size on metabolism, therefore, have broad implications for plant biogeography and for other theories of plant ecology and suggest that selection on metabolism may have a role in genome size evolution (Rodd et al. 2020). Therefore, it should be noted that, as seen by Parma et al. (*submitted*), among the analyzed Cleomaceae species, the plants with the largest genome are those that have a wide geographic distribution, in addition to having the C_4 photosynthetic metabolism or a clearly modified C_3 metabolism. In addition, there is a homoplastic tendency of some groups, suggesting that several plant lineages possess traits that increase to progress to the C_4 photosynthetic mechanism (Christin et al., 2010; Sage, 2001), as already verified for Cleomaceae (Bergh et al. 2014). Thus, such enablers may include genomic factors, as well as ecological and anatomical properties (Sage, 2001), as previously suggested (Parma et al. *submitted*; Marshall et al., 2007) and evidenced in the present study for Cleomaceae species.

Accordingly, based on the exposed and the results presented in this work, we suggest that *Tarenaya* sp., *T. longicarpa, T. aculeata, T. microcarpa* and *T. rosea* accessions (groups I-III) are in stage two or three of evolution for C_4 photosynthesis. These stages include (i) genome duplication; (ii) increased density of venation and, (iii) increased BSC (Bergh et al., 2014; Gowik and Westhoff, 2011; Sage, 2004). Accordingly, as observed to grasses (Christian et al., 2013), these results suggest that key alterations of foliar anatomy are occurring in a C_3 context and preceding the emergence of the C_4 syndrome. However, further analyses such as quantifying the number of organelles in BSC are still necessary to classify the species according to the stage of C_4 photosynthesis evolution of they are.

Is the Brazilian *G. gynandra* access a new C_4 species?

In this study, we demonstrated that the Brazilian access of *G. gynandra* exhibits a low CO_2 compensation point and a lower venation density, even lower than those observed for other *G. gynandra* accessions (*e.g.* Marshall et al., 2007; Voznesenskaya et al., 2007; Reeves et al.,

2018) (Fig. 2E). Interestingly, the leaves of this accession also have Kranz anatomy, with larger BS cells as well as shorter interveins distance (Sup. Fig. 4O). Together with the low CO_2 compensation point, these are typical traits of C_4 plants (Sage, 2004).

Noteworthy the Brazilian access of *G. gynandra* has some physiological and anatomical differences when compared to the accessions studied in previous studies (Marshall et al., 2007; Voznesenskaya et al., 2007, 2018; Reeves et al., 2018), although the experimental conditions were similar with our experimental approach. In this way, among the analyzed traits, we highlight the variables A_N, CO_2 compensation point, $\delta^{13}C$, WUEi and vein density (Table 2). The differences found here for *G. gynandra* compared to other studies with the species, leads us to believe that our access clearly has a genetic diversity, compared to the others, as identified in the work of Reeves et al. (2018) for accesses from Africa and Asia. Thus, when using *G. gynandra* ITS sequences from different continents (Africa and Asia; Reeves et al. (2018) and South America (present study)), we verified that the accessions were grouped in different branches (Sup. Fig. 5). Therefore, the *G. gynandra* from South America may be in the process of adaptation/evolution to differentiate itself from other accesses (Africa and Asia). In this sense, it is worth analyzing these accessions more deeply, mainly in relation to their morphology and ecology, since something similar happened with the species *T. hassleriana* (Brazil) and *T. werdermannii* (Bolivia). These species of *Tarenaya* have recently diversified by sympatric speciation (Soares-Neto et al., 2018), that is, when gene flow no longer occurs due to geographic isolation between accessions. Therefore, it is expected that soon it will be possible to recognize the different accessions of *G. gynandra* as a new taxon.

Table 2: Physiological variation for photosynthetic gas-exchange and anatomical parameters among diverse accessions of *Gynandropsis gynandra*.

CO_2 Γ (µmol mol^{-1})	***A_N*** (µmol m^{-2}s^{-1})	**$\delta^{13}C$ (‰)**	**Vein density (mm mm^{-2})**	**WUE (µmol mmol^{-1})**	**Reference**
9	-	-13.9	8	-	(Marshall et al., 2007)
6.8	~25	-15	-	6.9	(Voznesenskaya et al., 2007)
~2	25	-14.8	-	-	(Voznesenskaya et al., 2018)

2-6	20-30	-15.5-17	6-10	5-12	(Reeves et al., 2018)
-	25	-15	12	5-7	(Parma et al. *submitted*)
4.2	33	-	10	11	(Present study)

CONCLUSION

The Cleomaceae family exhibit a great morphological, physiological and anatomical diversity, being possible to observe this diversity in the group of studied species (15 spp.). In addition, it has been observed that some species (*e.g. Tarenaya* sp., *T. aculeata, T. microcarpa* and *T. rosea*) may be in transition/evolution to the C_3-C_4 intermediate photosynthetic mechanism. Moreover, the Brazilian access of *G. gynandra* used in this study has physiological, biochemical and anatomical characters coherent with C_4 species, but that differ from those found for other African and Asian accessions. Accordingly, these results indicate a genetic variation among the *G. gynandra* accessions worldwide. In addition, considering the great diversity of biochemical, physiological and anatomical traits found for *T. hassleriana* and *G. gynandra*, it is reasonable to assume that these accessions have a high genetic diversity and that these differences will help to elucidate important characteristics for the development of the C_4 photosynthetic mechanism.

ACKNOWLEDGMENT

This work was supported by funding from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; grant number 424024/2018-7), the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG; grant number APQ-00528-18 and CRA-RED00053-16). Research fellowships granted by CNPq to ANN and WLA as well as the scholarship granted by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to DFP are also gratefully acknowledged. MGMVV was supported by scholarships from CAPES (PNPD-1638006). APMW acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2048/1 – project ID 390686111.

REFERENCES

Anderson, J.T., Panetta, A.M., Mitchell-olds, T., 2012. Evolutionary and Ecological Responses to Anthropogenic Climate Change 1 160, 1728–1740. https://doi.org/10.1104/pp.112.206219

Bayat, S., Schranz, M.E., Roalson, E.H., Hall, J.C., 2018. Lessons from Cleomaceae, the Sister of Crucifers. Trends Plant Sci. 23, 808–821. https://doi.org/10.1016/j.tplants.2018.06.010

Bergh, E. Van Den, Külahoglu, C., Bräutigam, A., Hibberd, J.M., Weber, A.P.M., Zhu, X., Schranz, M.E., 2014. Current Plant Biology Gene and genome duplications and the origin of C 4 photosynthesis : Birth of a trait in the Cleomaceae. Curr. Plant Biol. 1, 2–9. https://doi.org/10.1016/j.cpb.2014.08.001

Bianconi, M.E., Dunning, L.T., Moreno-Villena, J.J., Osborne, C.P., Christin, P.A., 2018. Gene duplication and dosage effects during the early emergence of C 4 photosynthesis in the grass genus Alloteropsis. J. Exp. Bot. 69, 1967–1980. https://doi.org/10.1093/jxb/ery029

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Christin, P., Osborne, C.P., Chatelet, D.S., Columbus, J.T., Besnard, G., Hodkinson, T.R., Garrison, L.M., Vorontsova, M.S., Edwards, E.J., 2013. Anatomical enablers and the evolution of C 4 photosynthesis in grasses 110. https://doi.org/10.1073/pnas.1216777110

Cross, J.M., von Korff, M., Altmann, T., Bartzetko, L., Sulpice, R., Gibon, Y., Palacios, N., Stitt, M., 2006. Variation of enzyme activities and metabolite levels in 24 Arabidopsis accessions growing in carbon-limited conditions. Plant Physiol. 142, 1574–88. https://doi.org/10.1104/pp.106.086629

de Oliveira Silva, F.M., Lichtenstein, G., Alseekh, S., Rosado-Souza, L., Conte, M., Suguiyama, V.F., Lira, B.S., Fanourakis, D., Usadel, B., Bhering, L.L., DaMatta, F.M., Sulpice, R., Araújo, W.L., Rossi, M., de Setta, N., Fernie, A.R., Carrari, F., Nunes-Nesi, A., 2018a. The genetic architecture of photosynthesis and plant growth-related traits in tomato. Plant. Cell Environ. 41, 327–341. https://doi.org/10.1111/pce.13084

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Ethier, G.J., Livingston, N.J., 2004. On the need to incorporate sensitivity to CO2 transfer conductance into the Farquhar – von Caemmerer – Berry leaf.

Feodorova, T.A., Voznesenskaya, E. V, Edwards, G.E., Roalson, E.H. 2010. Biogeographic Patterns of Diversification and the Origins of C4 in Cleome (Cleomaceae). 35, 811–826. https://doi.org/10.1600/036364410X539880

Fernie, A.R., Roscher, A., Ratcliffe, R.G., Kruger, N.J., 2001. Fructose 2,6-bisphosphate activates pyrophosphate: fructose-6-phosphate 1-phosphotransferase and increases triose phosphate to hexose phosphate cycling in heterotrophic cells. Planta 212, 250–263. https://doi.org/10.1007/s004250000386

Flexas, J., Ortuño, M.F., Ribas-Carbo, M., Diaz-Espejo, A., Flórez-Sarasa, I.D., Medrano, H., 2007. Mesophyll conductance to CO2 in Arabidopsis thaliana. New Phytol. 175, 501–511. https://doi.org/10.1111/j.1469-8137.2007.02111.x

Flood, P.J., 2019. Using natural variation to understand the evolutionary pressures on plant photosynthesis. Curr. Opin. Plant Biol. https://doi.org/10.1016/j.pbi.2019.06.001

Genty, B., Briantais, J.M., Baker, N.R., 1989. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. Biochim. Biophys. Acta - Gen. Subj. 990, 87–92. https://doi.org/10.1016/S0304-4165(89)80016-9

Gowik, U., Bra, A., 2011. Evolution of C4 Photosynthesis in the Genus Flaveria : How Many and Which Genes Does It Take to Make C4 ? 23, 2087–2105.

https://doi.org/10.1105/tpc.111.086264

Gowik, U., Westhoff, P., 2011. The Path from C 3 to C 4 Photosynthesis 1 155, 56–63. https://doi.org/10.1104/pp.110.165308

Grassi, G., Magnani, F., 2005. Stomatal , mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees 834–849.

Harley, P.C., Loreto, F., Marco, G. Di, Sharkey, T.D., Scalo, M., Al, H.E.T., 1992. Theoretical Considerations when Estimating the Mesophyll Conductance to CO2 Flux by Analysis of the Response of Photosynthesis to CO21 1429–1436.

Heckmann, D., Schulze, S., Denton, A., Gowik, U., Westhoff, P., Weber, A.P.M., Lercher, M.J., 2013. Theory Predicting C4 Photosynthesis Evolution : Modular , Individually Adaptive Steps on a Mount Fuji Fitness Landscape. Cell 153, 1579–1588. https://doi.org/10.1016/j.cell.2013.04.058

Hervé, D., Fabre, F., Flores Berrios, E., Leroux, N., Al Chaarani, G., Planchon, C., Sarrafi, A., Gentzbittel, L., 2001. QTL analysis of photosynthesis and water status traits in sunflower (Helianthus annuus L.) under greenhouse conditions. J. Exp. Bot. 52, 1857–1864. https://doi.org/10.1093/jexbot/52.362.1857

Inda, L.A., Torrecilla, Æ.P., Catala, Æ.P., Ruiz-zapata, T., 2008. Phylogeny of Cleome L. and its close relatives Podandrogyne Ducke and Polanisia Raf. (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology 111–126. https://doi.org/10.1007/s00606-008-0026-y

International Plant Names Index [WWW Document], n.d. URL https://www.ipni.org/ (accessed 1.28.20).

Koch, M., Haubold, B., Mitchell-Olds, T., 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. Am. J. Bot. 88, 534–44.

Koteyeva, N.K., Voznesenskaya, E. V, Roalson, E.H., Edwards, G.E., 2011. Diversity in forms of C4 in the genus Cleome (Cleomaceae) 269–283. https://doi.org/10.1093/aob/mcq239

Lawson, T., Kramer, D.M., Raines, C.A., 2012. Improving yield by exploiting mechanisms underlying natural variation of photosynthesis. Curr. Opin. Biotechnol. https://doi.org/10.1016/j.copbio.2011.12.012

Logan, B.A., Adams, W.W., Demmig-Adams, B., 2007. Viewpoint: Avoiding common pitfalls of chlorophyll fluorescence analysis under field conditions. Funct. Plant Biol. 34, 853. https://doi.org/10.1071/FP07113

Luciano, R., Neto, S., Thomas, W.W., Regina, M., Barbosa, D.V., Roalson, E.H., 2018. New combinations and taxonomic notes for Tarenaya (Cleomaceae) 32, 540–545. https://doi.org/10.1590/0102-33062017abb0417

Marshall, D.M., Muhaidat, R., Brown, N.J., Liu, Z., Stanley, S., Griffiths, H., Sage, R.F., Hibberd, J.M., 2007. Cleome , a genus closely related to Arabidopsis , contains species spanning a developmental progression from C3 to C4 photosynthesis 886–896. https://doi.org/10.1111/j.1365-313X.2007.03188.x

Mckown, A.D., Dengler, N.G., 2007. Key innovations in the evolution of Kranz anatomy and C4 vein pattern in Flaveria (Asteraceae). Am. J. Bot. 94, 382–389. https://doi.org/10.3732/ajb.94.3.382

Monson, R.K., 2003. Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. Int. J. Plant Sci. https://doi.org/10.1086/368400

Muhaidat, R., Sage, T.L., Frohlich, M.W., Dengler, N.G., Sage, R.F., 2011. Characterization of C3-C4 intermediate species in the genus *Heliotropium* L. (Boraginaceae): Anatomy, ultrastructure and enzyme activity. Plant, Cell Environ. 34, 1723–1736. https://doi.org/10.1111/j.1365-3040.2011.02367.x

Nunes-Nesi, A., Carrari, F., Gibon, Y., Sulpice, R., Lytovchenko, A., Fisahn, J., Graham, J., Ratcliffe, R.G., Sweetlove, L.J., Fernie, A.R., 2007. Deficiency of mitochondrial fumarase activity in tomato plants impairs photosynthesis via an effect on stomatal function. Plant J. 50, 1093–1106. https://doi.org/10.1111/j.1365-313X.2007.03115.x

Nylander, 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

O'Brien, T.P., Feder, N., McCully, M.E., 1964. Polychromatic staining of plant cell walls by toluidine blue O. Protoplasma 59, 368–373. https://doi.org/10.1007/BF01248568

Olmstead, R.G., Jennifer, A., Wolfe, K.H., Biology, E.P.O., 1993. Update section Ninety extra nucleotide in ndhF gene of tobacco chloroplast DNA : a summary of revisions to the 1986 genome sequence 1191–1193.

Osborne, C.P., Sack, L., 2012. Evolution of C4 plants: A new hypothesis for an interaction of CO2 and water relations mediated by plant hydraulics. Philos. Trans. R. Soc. B Biol. Sci. 367, 583–600. https://doi.org/10.1098/rstb.2011.0261

Patchell, M.J., Roalson, E.H., Hall, J.C., 2014. Resolved phylogeny of Cleomaceae based on all three genomes 63, 315–328.

Pons, T.L., Flexas, J., Caemmerer, S. Von, Evans, J.R., Genty, B., Ribas-carbo, M., Brugnoli, E., 2009. Estimating mesophyll conductance to CO 2 : methodology , potential errors , and recommendations 60, 2217–2234. https://doi.org/10.1093/jxb/erp081

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Syst. Biol. https://doi.org/10.1080/10635150490522304

Reeves, G., Singh, P., Rossberg, T.A., Sogbohossou, E.O.D., Schranz, M.E., Hibberd, J.M., 2018. Natural variation within a species for traits underpinning C 4 photosynthesis. Plant Physiol. 177, 504–512. https://doi.org/10.1104/pp.18.00168

Roddy, A.B., Théroux-Rancourt, G., Abbo, T., Benedetti, J.W., Brodersen, C.R., Castro, M., Castro, S., Gilbride, A.B., Jensen, B., Jiang, G.F., Perkins, J.A., Perkins, S.D., Loureiro, J., Syed, Z., Alexander Thompson, R., Kuebbing, S.E., Simonin, K.A., 2020. The scaling of genome size and cell size limits maximum rates of photosynthesis with implications for ecological strategies. Int. J. Plant Sci. 181, 75–87. https://doi.org/10.1086/706186

Rodeghiero, M., Niinemets, Ü., Cescatti, A., 2007. Major diffusion leaks of clamp-on leaf cuvettes still unaccounted: How erroneous are the estimates of Farquhar et al. model parameters? Plant, Cell Environ. 30, 1006–1022. https://doi.org/10.1111/j.1365-3040.2007.001689.x

Rosado-Souza, L., Scossa, F., Chaves, I.S., Kleessen, S., Salvador, L.F.D., Milagre, J.C., Finger, F., Bhering, L.L., Sulpice, R., Araújo, W.L., Nikoloski, Z., Fernie, A.R., Nunes-Nesi, A., 2015. Exploring natural variation of photosynthetic, primary metabolism and growth parameters in a large panel of Capsicum chinense accessions. Planta 242, 677–691. https://doi.org/10.1007/s00425-015-2332-2

Sage, R.F., 2004. The evolution of C4 photosynthesis. New Phytol. https://doi.org/10.1111/j.1469-8137.2004.00974.x

Sage, R.F., 2001. Environmental and evolutionary preconditions for the origin and diversification of the C4 photosynthetic syndrome. Plant Biol. https://doi.org/10.1055/s-2001-15206

Sage, R.F., Sage, R.F., 2004. The evolution of C 4 photosynthesis 341–370. https://doi.org/10.1046/j.1469-8137.2004.00974.x

Schlüter, U., Bräutigam, A., Gowik, U., Melzer, M., Christin, P., Kurz, S., Mettler-altmann, T., Weber, A.P.M., 2017. Photosynthesis in C 3 – C 4 intermediate Moricandia species 68, 191–206. https://doi.org/10.1093/jxb/erw391

Schuler, M.L., Mantegazza, O., Weber, A.P.M., 2016. Engineering C 4 photosynthesis into C

3 chassis in the synthetic biology age 51–65. https://doi.org/10.1111/tpj.13155

Schulze, S., Mallmann, J., Burscheidt, J., Koczor, M., Streubel, M., Bauwe, H., Gowik, U., Westhoff, P., 2013. Evolution of C 4 Photosynthesis in the Genus Flaveria : Establishment of a Photorespiratory CO 2 Pump 25, 2522–2535. https://doi.org/10.1105/tpc.113.114520

Sharkey, T.D., Bernacchi, C.J., Farquhar, G.D., Singsaas, E.L., 2007. Fitting photosynthetic carbon dioxide response curves for C3 leaves. Plant, Cell Environ. 30, 1035–1040. https://doi.org/10.1111/j.1365-3040.2007.01710.x

Sogbohossou, E.O.D., Kortekaas, D., Achigan-Dako, E.G., Maundu, P., Stoilova, T., Van Deynze, A., de Vos, R.C.H., Schranz, M.E., 2019. Association between vitamin content, plant morphology and geographical origin in a worldwide collection of the orphan crop Gynandropsis gynandra (Cleomaceae). Planta 250, 933–947. https://doi.org/10.1007/s00425-019-03142-1

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

U, S., AP, W., 2016. The Road to C4 Photosynthesis: Evolution of a Complex Trait via Intermediary States. Plant Cell Physiol. 57. https://doi.org/10.1093/PCP/PCW009

van Bezouw, R.F.H.M., Keurentjes, J.J.B., Harbinson, J., Aarts, M.G.M., 2019. Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency. Plant J. 97, 112–133. https://doi.org/10.1111/tpj.14190

Voznesenskaya, E. V., Koteyeva, N.K., Chuong, S.D.X., Ivanova, A.N., Barroca, J., Craven, L.A., Edwards, G.E., 2007. Physiological, anatomical and biochemical characterisation of photosynthetic types in genus Cleome (Cleomaceae). Funct. Plant Biol. 34, 247. https://doi.org/10.1071/FP06287

Walker, B.J., Cousins, A.B., 2013. Influence of temperature on measurements of the CO 2 compensation point : differences between the Laisk and O 2 - exchange methods 64, 1893–1905. https://doi.org/10.1093/jxb/ert058

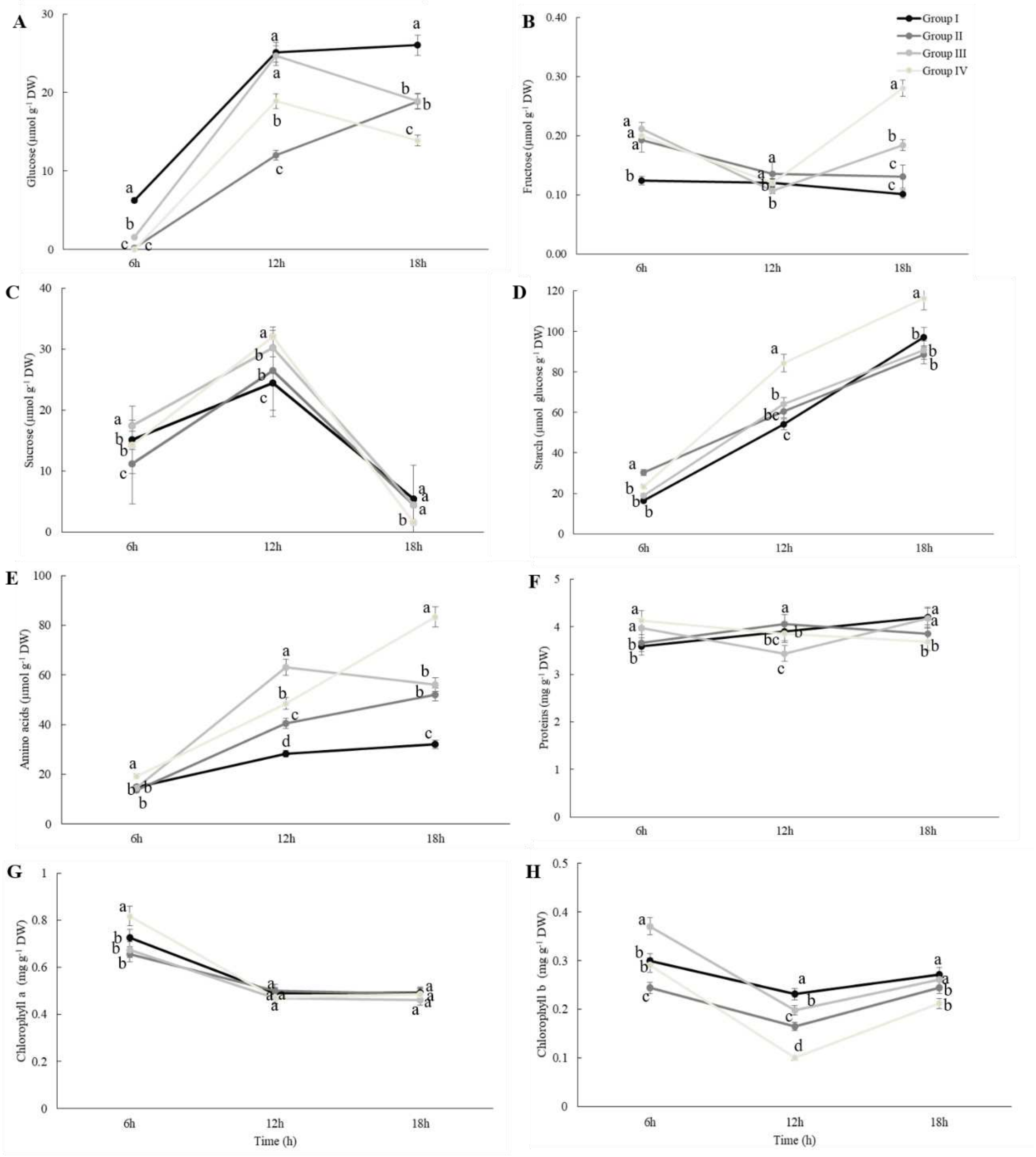
Wang, X., Gowik, U., Tang, H., Bowers, J.E., Westhoff, P., Paterson, A.H., 2009. Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. Genome Biol. 10, R68. https://doi.org/10.1186/gb-2009-10-6-r68

Williams, B.P., Aubry, S., Hibberd, J.M., 2012. Molecular evolution of genes recruited into C 4 photosynthesis. Trends Plant Sci. https://doi.org/10.1016/j.tplants.2012.01.008

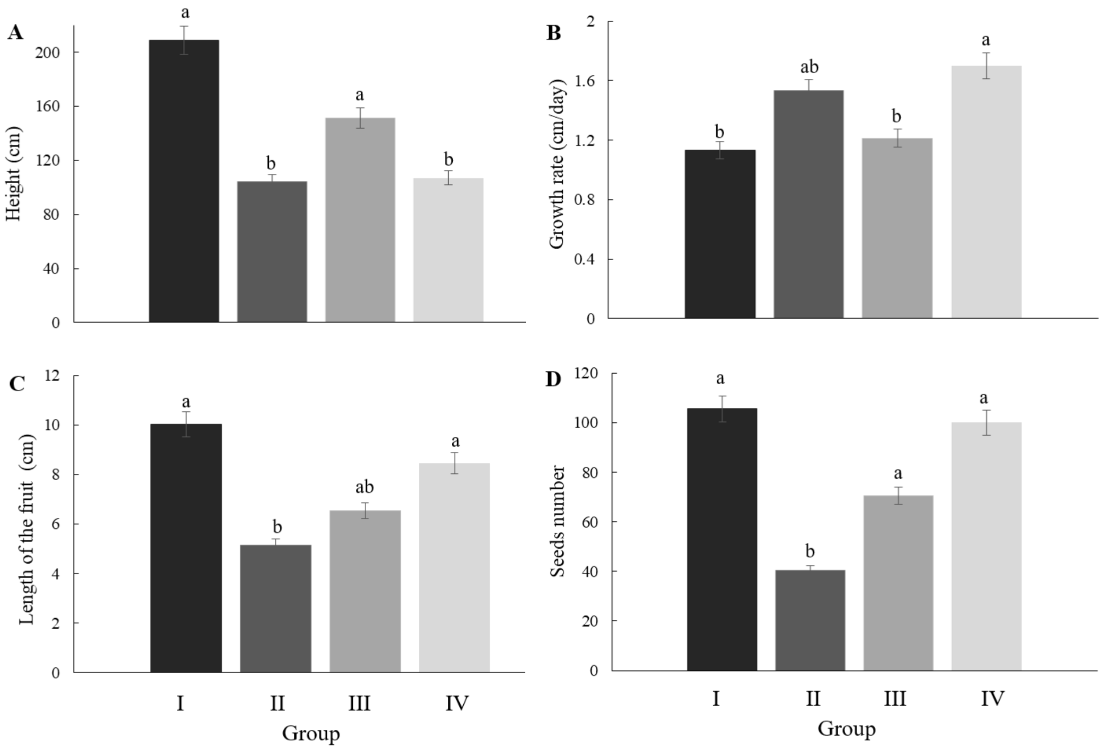
Yin, X., Struik, P.C., Romero, P., Harbinson, J., Evers, J.B., Van Der Putten, P.E.L., Vos, J., 2009. Using combined measurements of gas exchange and chlorophyll fluorescence to estimate parameters of a biochemical C3 photosynthesis model: A critical appraisal and a new integrated approach applied to leaves in a wheat (Triticum aestivum) canopy. Plant, Cell Environ. 32, 448–464. https://doi.org/10.1111/j.1365-3040.2009.01934.x

Zsögön, A., Alves Negrini, A.C., Peres, L.E.P., Nguyen, H.T., Ball, M.C., 2015. A mutation that eliminates bundle sheath extensions reduces leaf hydraulic conductance, stomatal conductance and assimilation rates in tomato (*Solanum lycopersicum*). New Phytol. 205, 618–626. https://doi.org/10.1111/nph.13084.

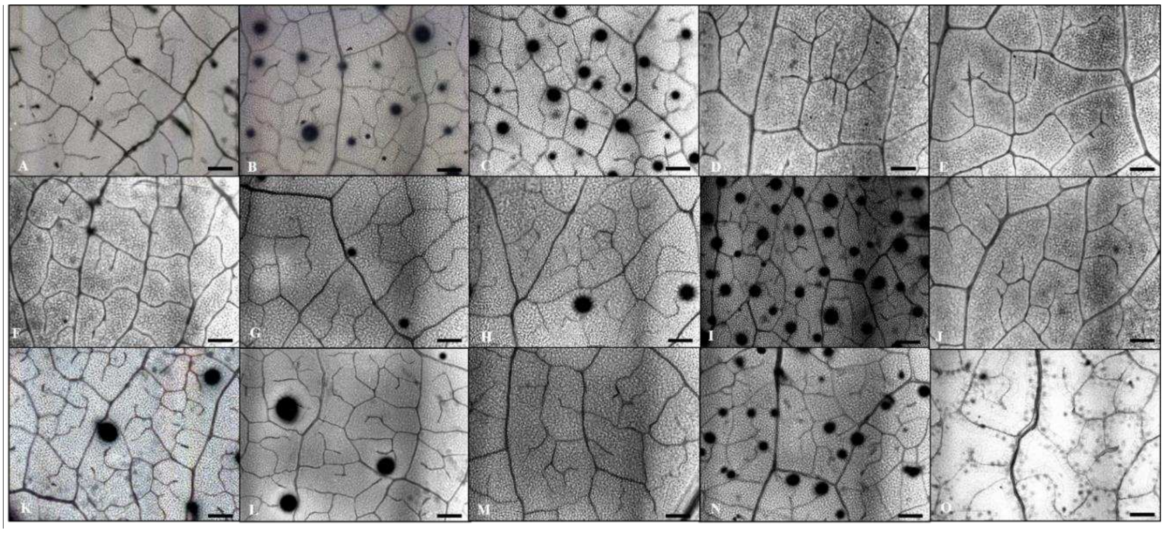
SUPPLEMENTARY MATERIAL



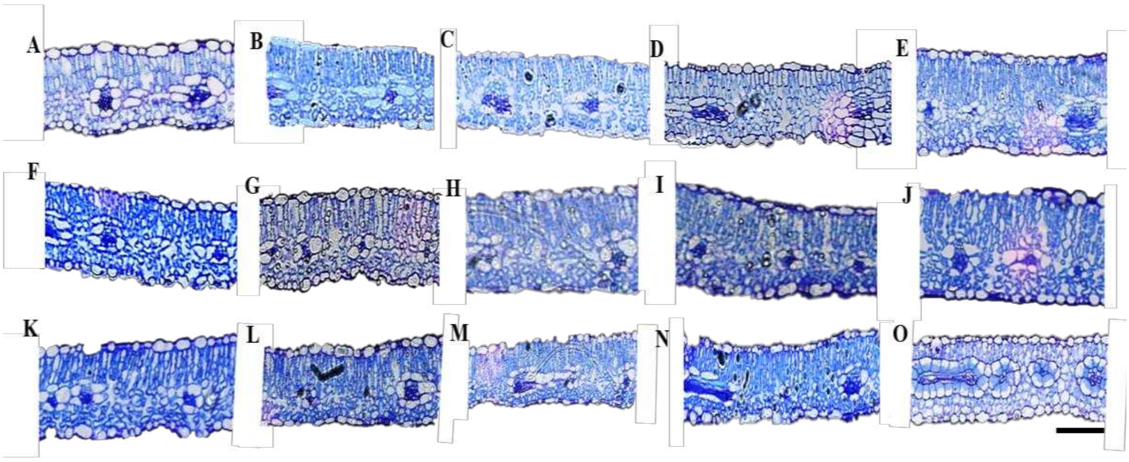
Sup. Figure 1. Levels of metabolites extracted from leaves of a subset of Cleomaceae species. For all metabolic analysis, leaf samples were harvested at beginning, middle and end of the light period. Fully expanded leaves from the third node were used. The plants were in a vegetative stage; groups I and III (shrub species) were four months old and groups II and IV (herbaceous species) were two months old. A- sucrose. B- fructose. C- glucose. D- starch. E- total protein. F- total free amino acids. G- chlorophyll *a*. H- chlorophyll *b*. Letters above individual spots indicate significant groupings according to Tukey's Test, n=5.



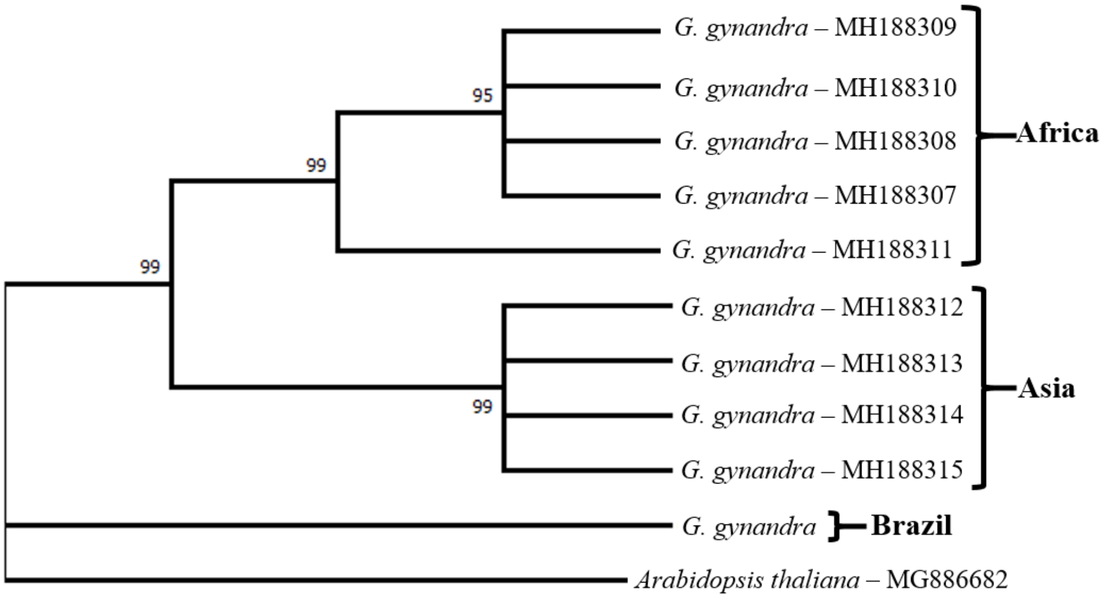
Sup. Figure 2. Morphological characteristics. A- height, B- growth rate, C- length of the fruit. D- seeds number per fruit. Letters above individual bars indicate significant groupings according to Tukey's Test, n=5.



Sup. Figure 3. Vein density in leaves of Cleomaceae morphospecies. Diaphanization of the fully expanded leaf of the third node. A-C, G-N with four months of age and D-F, O with two months. Spots indicate presence of secretory trichomes. A- TAM: *Tarenaya* sp. (Manaus-AM); B- TS: *T. spinosa* (Teresina-PI); C- TL: *T. longicarpa* (Picos-PI); D-TM: *T. microcarpa* (Belém-PA); E- TD: *T. diffusa* (Feira de Santana-BA); F- TA: *T. aculeata* (Feira de Santana-BA); G- THV: *T. hassleriana* (Viçosa-MG); H-THS: *T. hassleriana* (São Miguel-MG); I- TR: *T. rosea* (Colatina-ES); J- CP: *C. paludosa* (Belém-PA); K- THC: *T. hassleriana* (Canaã-MG); L- THD: *T. hassleriana* (Domingos Martins-ES); M- THJ: *T. hassleriana* (Joinville-SC); N- TP: *T. parviflora* (Pombal-PB); O- GG: *G. gynandra* (Mossoró-RN). Species name (City of Sampling-state). 10x. Bars = 10 µm.



Sup. Figure 4. Leaves cross section of the fully expanded leaf of the third node of Cleomaceae morphospecies. A-C, G-N with four months of age and D-F, O with two months. A- TAM: *Tarenaya* sp. (Manaus-AM); B- TS: *T. spinosa* (Teresina-PI); C- TL: *T. longicarpa* (Picos-PI); D-TM: *T. microcarpa* (Belém-PA); E- TD: *T. diffusa* (Feira de Santana-BA); F- TA: *T. aculeata* (Feira de Santana-BA); G- THV: *T. hassleriana* (Viçosa-MG); H-THS: *T. hassleriana* (São Miguel-MG); I- TR: *T. rosea* (Colatina-ES); J- CP: *C. paludosa* (Belém-PA); K- THC: *T. hassleriana* (Canaã-MG); L- THD: *T. hassleriana* (Domingos Martins-ES); M- THJ: *T. hassleriana* (Joinville-SC); N- TP: *T. parviflora* (Pombal-PB); O- GG: *G. gynandra* (Mossoró-RN). Species name (City of Sampling-state). Bars = 100 µm.



Sup. Figure 5. Phylogenetic relationships among *Gynandropsis gynandra* accessions by Maximum Likelihood inference sequences of nuclear ITS. Numbers at nodes indicate bootstrap values.

Sup. Table 1A. Photosynthetic characterization of Cleomaceae species. Values are presented as means ± SE (n=5) obtained using the third leaf totally expanded from five different plants per species. Letters indicate significant groupings according to Tukey's Test, n=5.

Group	Species	A_N (µmol CO_2 m^{-2} s^{1})	A_{GROSS} (µmol CO_2 m^{-2} s^{-1})	g_s (mol H_2O m^{-2} s^{-1})	E (mmol H_2O m^{-2} s^{1})	Γ (µmol mol^{-1})	WUEi (A_N/g_s)	WUE (A_N/E)	R_d (µmol CO_2 m^{-2} s^{-1})	Pr (µmol CO_2 m^{-2} s^{-1})	F_v/F_m	F_v'/F_m'
I	TAM	28.87 ± 0.78a	41.84 ± 0.46a	0.91 ± 0.044a	11.33 ± 0.47b	62.19 ± 1.17b	31.97 ± 2.04f	2.57 ± 0.15e	3.31 ± 0.01b	9.66 ± 0.54c	0.80 ± 0.008a	0.71 ± 0.005a
	TL	29.05 ± 2.20a	41.13 ± 0.44a	0.87 ± 0.047b	11.98 ± 0.91b	78.98 ± 1.85^{a}	33.10 ± 0.70f	2.42 ± 0.01e	2.93 ± 0.01c	9.15 ± 1.75c	0.79 ± 0.000a	0.70 ± 0.007a
	TS	28.41 ± 0.79a	41.45 ± 0.59a	0.81 ± 0.118b	10.05 ± 1.43b	68.08 ± 1.15b	41.06 ± 10.16e	3.33 ± 0.84d	4.03 ± 0.01a	9.01 ±0.46c	0.79 ± 0.000a	0.71 ± 0.004a
II	TA	25.22 ± 0.30b	36.11 ± 0.20b	0.29 ± 0.000g	4.72 ± 0.00d	55.07 ± 0.98c	84.81 ± 1.02b	5.34 ± 0.01b	4.93 ± 0.01a	6.17 ± 0.33g	0.73 ± 0.000b	0.59 ± 0.001c
	TD	17.26 ± 0.32d	30.42 ± 1.76c	0.29 ± 0.012g	4.00 ± 0.07d	62.26 ± 1.23b	58.04 ± 1.34d	4.31 ± 0.01c	4.93 ± 0.01a	8.22 ± 1.44e	0.78 ± 0.000a	0.61 ± 0.019c
	TM	25.22 ± 0.30b	42.46 ± 0.30a	0.29 ± 0.000g	4.72 ± 0.00d	52.08 ± 1.74c	84.81 ± 1.01b	5.34 ± 0.01b	4.99 ± 0.01a	12.25 ± 0.12a	0.79 ± 0.000a	0.70 ± 0.001a
III	CP	17.90 ± 4.39d	31.85 ± 4.32c	0.36 ± 0.002f	5.18 ± 0.08d	63.88 ± 1.25b	48.72 ± 12.21e	3.44 ± 0.79d	4.93 ± 0.01a	9.02 ± 0.65c	0.79 ± 0.002a	0.64 ± 0.036b
	THC	25.15 ± 0.01b	35.19 ± 1.58b	0.34 ± 0.000f	5.35 ± 0.00d	59.68 ± 1.52c	72.07 ± 1.02c	4.69 ± 0.01c	2.62 ± 0.01c	7.42 ± 1.58f	0.79 ± 0.002a	0.71 ± 0.001a
	THDM	18.27 ± 0.30d	30.56 ± 0.20c	0.40 ± 0.000e	5.51 ± 0.00d	62.72 ± 1.65b	44.71 ± 1.00e	3.31 ± 0.01d	2.94 ± 0.01c	9.34 ± 0.21c	0.79 ± 0.000a	0.65 ± 0.001b
	THJ	19.71 ± 1.35d	32.14 ± 0.63c	0.65 ± 0.083c	10.46 ± 0.68b	78.85 ± 1.22a	31.45 ± 3.75f	1.88 ±0.07f	3.08 ± 0.01b	9.34 ±1.19c	0.70 ± 0.045b	0.66 ± 0.015b
	THS	22.45 ± 1.02c	36.75 ± 0.90b	0.63 ± 0.068c	15.01 ± 1.28a	57.63 ± 0.95c	35.62 ± 2.44f	1.52 ± 0.19f	3.93 ± 0.01b	10.36 ± 0.67b	0.80 ± 0.000a	0.66 ± 0.018b
	THV	21.81 ± 0.77c	34.02 ± 1.01b	0.52 ± 0.063d	9.37 ± 1.76c	84.84 ± 2.63a	43.23 ± 3.72e	2.72 ± 0.56e	3.4 ± 0.01b	8.81 ± 0.33d	0.76 ± 0.014b	0.68 ± 0.007a
	TP	21.28 ± 0.92c	33.09 ± 1.09c	0.49 ±0.066e	7.25 ± 1.57c	65.57 ± 2.63b	44.71 ± 4.56e	3.27 ± 0.43d	3.43 ± 0.01b	8.38 ±0.28d	0.83 ± 0.000a	0.64 ± 0.011b
	TR	18.88 ± 3.58d	31.96 ± 3.64c	0.42 ± 0.150e	6.95 ± 1.69c	53.98 ± 0.94c	56.22 ± 7.62d	2.89 ± 0.19e	4.45 ± 0.01a	8.62 ± 0.28d	0.79 ± 0.000a	0.64 ± 0.023b
IV	GG	33.81 ± 0.31a	35.26 ± 0.33b	0.12 ± 0.000h	2.24 ± 0.00e	4.22 ± 0.02d	203.34 ± 2.01a	11.07 ± 0.01a	4.27 ± 0.01a	6.17 ± 0.33g	0.73 ± 0.000b	0.59 ± 0.001c

Cleomaceae morphospecies. A- TAM: *Tarenaya* sp. (Manaus-AM); B- TS: *T. spinosa* (Teresina-PI); C- TL: *T. longicarpa* (Picos-PI); D-TM: *T. microcarpa* (Belém-PA); E- TD: *T. diffusa* (Feira de Santana-BA); F- TA: *T. aculeata* (Feira de Santana-BA); G- THV: *T. hassleriana* (Viçosa-MG); H-THS: *T. hassleriana* (São Miguel-MG); I- TR: *T. rosea* (Colatina-ES); J- CP: *C. paludosa* (Belém-PA); K- THC: *T. hassleriana* (Canaã-MG); L- THD: *T. hassleriana* (Domingos Martins-ES); M- THJ: *T. hassleriana* (Joinville-SC); N- TP: *T. parviflora* (Pombal-PB); O- GG: *G. gynandra* (Mossoró-RN). Species name (City of Sampling-state).

Sup. Table 1B. Photosynthetic characterization of Cleomaceae species. Values are presented as means ± SE (n=5) obtained using the third leaf totally expanded from five different plants per species. Letters indicate significant groupings according to Tukey's Test, n=5.

Group	Species	J_{flu} (µmol m^{-2} s^{-1})	C_i (µmol CO_2 m^{-1})	g_m (mol CO_2 m^{-2} s^{-1} bar^{-1})	V_{cmax_Ci} (µmol $m^{-2}s^{-1}$)	V_{cmax_Cc} (µmol $m^{-2}s^{-1}$)	J_{max_Ci} (µmol $m^{-2}s^{-1}$)	J_{max_Cc} (µmol $m^{-2}s^{-1}$)	Stomatal limitation	Mesophyll limitation	Biochemical limitation
I	TAM	224.00 ± 6.24a	313.93 ± 3.61a	0.04 ± 0.003b	93.97 ± 3.19b	24.71 ± 3.12d	144.85 ± 25.77a	124.40 ± 9.70a	0.01 ± 0.000c	0.45 ± 0.036a	0.54 ± 0.036e
	TL	217.74 ± 14.87a	311.27 ± 3.51a	0.03 ± 0.001c	23.14 ± 1.63f	25.44 ± 2.98d	46.73 ± 1.15d	49.71 ± 2.80d	0.01 ± 0.000c	0.26 ± 0.010b	0.73 ± 0.010c
	TS	227.23 ± 6.73a	300.31 ± 15.37a	0.05 ± 0.001b	29.00 ± 1.52e	28.17 ± 1.59d	53.42 ± 2.14d	52.13 ± 1.56d	0.001 ± 0.003d	0.40 ± 0.003a	0.58 ± 0.059e
II	TA	197.21 ± 6.42b	197.21 ± 6.45c	0.07 ± 0.001a	98.79 ± 1.00a	8.36 ± 0.70f	173.51 ± 10.80a	93.32 ± 1.30b	0.02 ± 0.000c	0.18 ± 0.002c	0.79 ± 0.002b
	TD	227.69 ± 6.01a	283.31 ± 1.76b	0.03 ± 0.005c	51.70 ± 2.56c	17.96 ± 3.47e	94.32 ± 10.08c	99.91 ± 26.12b	0.01 ± 0.000c	0.09 ± 0.019d	0.89 ± 0.019a
	TM	227.69 ± 5.63a	234.11 ± 1.00a	0.01 ± 0.001d	98.79 ± 1.00a	22.22 ± 1.60d	154.53 ± 8.30a	138.49 ± 15.70a	0.02 ± 0.000c	0.17 ± 0.010c	0.80 ± 0.010b
III	CP	166.44 ± 37.09c	298.26 ± 22.67b	0.05 ± 0.015b	54.85 ± 12.38c	33.64 ± 9.68c	123.90 ± 26.25b	120.10 ± 22.85a	0.06 ± 0.018b	0.23 ± 0.011c	0.70 ± 0.030c
	THC	214.03 ± 10.09b	254.42 ± 1.00b	0.07 ± 0.006a	89.06 ± 0.10b	42.00 ± 2.68b	164.45 ± 7.00a	131.44 ± 10.62a	0.06 ± 0.005b	0.18 ± 0.002c	0.75 ± 0.005c
	THDM	150.58 ± 5.00d	304.04 ± 2.77a	0.04 ± 0.001c	58.21 ± 0.01c	9.28 ± 0.58f	131.00 ± 1.10b	72.92 ± 10.00c	0.02 ± 0.000c	0.12 ± 0.010d	0.84 ± 0.010b
	THJ	204.94 ± 15.73b	320.24 ± 7.64a	0.03 ± 0.002c	31.04 ± 1.95e	13.63 ± 3.21e	89.69 ± 1.33c	84.02 ± 8.43b	0.01 ± 0.002c	0.07 ± 0.015e	0.92 ± 0.017a
	THS	223.39 ± 10.89a	306.01 ± 5.33a	0.06 ± 0.007a	29.56 ± 1.68e	21.45 ± 9.25d	97.93 ± 6.97c	105.82 ± 15.69b	0.01 ± 0.000c	0.05 ± 0.004e	0.94 ± 0.004a
	THV	199.82 ± 6.41b	299.77 ± 4.58b	0.50 ± 0.150b	54.69 ± 7.27c	79.33 ± 27.61a	142.27 ± 1.10a	144.04 ± 19.82a	0.12 ± 0.024a	0.11 ± 0.039d	0.80 ± 0.057b
	TP	184.34 ± 6.79c	300.33 ± 6.28a	0.02 ± 0.001c	40.77 ± 11.75d	18.82 ± 2.22e	43.82 ± 6.98d	51.90 ± 4.82d	0.01 ± 0.000c	0.32 ± 0.064b	0.67 ± 0.065d
	TR	192.61 ± 9.90b	281.86 ± 9.90b	0.01 ± 0.003d	24.85 ± 0.87f	25.52 ± 0.62d	48.20 ± 1.10d	49.97 ± 0.97d	0.01 ± 0.004c	0.26 ± 0.070b	0.72 ± 0.068c
IV	GG	197.21 ± 6.42b	197.21 ± 6.45c	0.07 ± 0.001a	103.73 ± 1.63a	30.93 ± 1.60c	182.20 ± 6.90a	143.99 ± 12.00a	0.02 ± 0.005c	0.22 ± 0.000c	0.75 ± 0.005c

Cleomaceae morphospecies as described in the last table (Sup. Table 1A).

CHAPTER 5

EXPLORING THE FLORAL DIVERSITY OF CLEOMACEAE SPECIES: VALUE OF FLORAL DIVERSITY DURING POLLINATOR CRISIS

Daniele F. Parma[1], Kaik F. Souza[1], Marcelo G. M. V. Vaz[1], Wagner L. Araújo[1], Agustin Zsögön[1], Andreas P.M. Weber[2], Eric Schranz[3] & Adriano Nunes-Nesi[1]*

[1] Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil

[2] Institute for Plant Biochemistry Biology, Cluster of Excellence on Plant Science CEPLAS, Heinrich Heine University, Düsseldorf, Germany

[3] Biosystematics Group, Wageningen University, Droevendaalsesteeg, Wageningen, The Netherlands

*Corresponding author:
Adriano Nunes-Nesi
Departamento de Biologia Vegetal,
Universidade Federal de Viçosa,
36570-900 Viçosa, Minas Gerais, Brazil
Phone: +55-31-3612-5357
Email: nunesnesi@ufv.br

Background and Aims The angiosperm family Cleomaceae has been studied primarily in terms of its transitional C_3–C_4 metabolism and floral characteristics. The family is characterized by remarkable floral diversity and abundant nectar and pollen production. In this context, the present study evaluated the morphoanatomical characteristics associated with the floral diversity of 10 Cleomaceae species, in addition to respectively 7 and 5 Brazilian accessions of *Tarenaya hassleriana* and *T. longicarpa*.

Methods We examined the flowers (all morphs) produced through five types of crosses and determined the most efficient type based on the number of seeds produced. In addition, we evaluated the pollen limitation, self-incompatibility index, and self-pollination index and explored the possible effects of pollination type on seed germination.

Key Results The studied species formed three groups based on sexual expression: andromonoecious, hermaphrodite (chasmogamous and cleistogamous flowers, being respectively the last and first record for the family) and polygamous flowers. Heterostyly and dichogamy were also observed. Notably, the studied species were self-compatible, without pollen limitation. *Gynandropsis gynandra*, *T. aculeata*, *T. diffusa*, and *T. microcarpa* exhibited autogamy, whereas the rest of the species did not undergo spontaneous self-pollination.

Conclusions Overall, hand cross-pollination produced greater seed yield, followed by open pollination. Natural self-pollination produced the lowest number of seeds (n = 18), with no seed production noted in 16 accessions. Interestingly, the pollination type did not affect seed germination. Nectar and pollen production can be explored as the attributes related to the attraction and conservation of insects during pollinator crisis.

Keywords: Andromonoecy, Cleomaceae, *Cleoserrata,* dichogamy, *Gynandropsis*, hercogamy, *Tarenaya*

INTRODUCTION

One of the great mysteries in plant biology is the extraordinary diversity of angiosperm species (Soltis *et al.*, 2019). Pollination is a mutually beneficial interaction between plants and pollinators (Marshman *et al.*, 2019). Thus, angiosperm species are likely to rapidly radiate because of their coevolution with specific insect pollinators (Ollerton *et al.*, 2011; Salzman *et al.*, 2021). Although crop domestication has entailed major changes in the plant reproductive systems, at least 75% of the species grown for human consumption depend on insects for pollination (Klein *et al.*, 2007; Kremen, 2018). This ecosystem service is strongly affected by nectar (carbohydrates), pollen (proteins), and other rewards (Prasifka *et al.,* 2018). For instance, pollination by bees plays critical roles in food security (Bailes *et al.*, 2015; Burkle *et al.*, 2017). Some of the most important crops, such as fruit trees of the Rosaceae family (apple, pear, plum, cherry, and almond), depend on insect pollinators (Ghazoul, 2005).

The global decline in insect pollinators is concerning, with increasing evidence of the associated threat to plant pollination (Hallmann *et al.*, 2017; Marshman *et al.*, 2019; Novais *et al.*, 2016). Anthropogenic environmental changes have promoted a parallel decline in both insect-pollinated plants and pollinators (Biesmeijer *et al.*, 2006). In general, the "pollinator crisis" is caused by habitat loss (mainly due to agricultural regions and other human-modified landscapes), climate change, and pesticide use (Ghazoul, 2005; Potts *et al.*, 2010; Marshman *et al.*, 2019; Heller *et al.*, 2019). Climate change, specifically global warming and water stress, may indirectly impact pollinators through their detrimental effects on the quantity or quality of floral resources (flowers, nectar, and pollen) (Mesgaran *et al.*, 2021), as verified in *Borago officinalis*, for instance (Descamps *et al.*, 2018). Pesticides, in turn, are fundamental to modern agriculture, and without their use, optimal productivity cannot be ensured (Popp *et al.*, 2013). Therefore, pesticides and herbicides are essential for preventing crop diseases as well as insect and weed infestation, which otherwise lead to yield losses (Popp *et al.*, 2013). Simultaneously, however, these benefits are detrimental to the natural ecosystems given their potential health risks. Pollution from agricultural intensification has decreased wildlife diversity (Cresswell *et al.*, 2018). For instance, the spraying of the insecticide fenitrothion in vineyards in northeast Italy lea to a decline in pollinator populations (Brittain *et al.*, 2010). In the face of this global challenge, knowledge of the genetic bases of floral structures is paramount, which would allow complete independence from external pollinating agents.

Cleomaceae, a sister family of Brassicaceae, has been studied for its transitional C_3–C_4 and C_4 photosynthetic mechanisms (Koteyeva et al., 2011; Marshall et al., 2007; Reeves et al., 2018; Parma et al. *submitted*). The family is distributed worldwide, comprising 26 genera and 270 species (Bayat et al., 2018; Stevens, 2001; Soares-Neto et al., 2020). Cleomaceae can be easily distinguished from other plant families based on specific morphological traits, such as compound leaves (3–12 leaflets), bracteate inflorescence, zygomorphic flowers, six stamens, and dry fruits with longitudinal dehiscence (Stevens, 2001; Iltis et al. 2011). Members of this family are of great economic and ecological importance, in addition to their use as model organisms in studies of floral evolution, C_4 photosynthesis, and comparative genomics and transcriptomics (Brown et al., 2005; Feodorova et al., 2010; Nozzolillo et al., 2010). Close relatives of Cleomaceae present high variation in floral morphs and associated features (e.g., nectar and pollen). As such, while some species harbor only hermaphrodite flowers (*Cleome violacea*; Cane, 2008), others bear male and hermaphrodite flowers (*Gynandropsis gynandra*; Omondi et al., 2017; Zohoungbogbo et al., 2018) and some even possess all floral morphs within the same individual (*Tarenaya spinosa*; Machado et al., 2006). Hermaphrodite Cleomaceae species display a predominantly self-pollinated reproductive pathway (autogamous with cleistogamous and chasmogamous flowers, *e.g.*, *Tarenaya aculeata* and *T. diffusa*), which may offer them an adaptive advantage over other cross-pollinated members that depend on external pollinators. Meanwhile, this reproductive mechanism has evolved at the cost of the loss of genetic variability, resulting in low tolerance of environmental changes, likely representing an evolutionary dead end (Guzella *et al.*, 2017; Takebayashi and Morrell, 2001).

In Brazil, 7 Cleomaceae genera (*Tarenaya*, *Gynandropsis*, *Melidiscus*, *Cleoserrata*, *Dactylaena*, *Haptocarpum*, and *Physostemon*) comprising 34 species have been described (Flora of Brazil, 2021). Among these, *Tarenaya* is the most diverse genus, with 22 species, and 10 of these are endemic (Flora of Brazil, 2021). Despite their wide distribution throughout the Brazilian territory, majority of the Cleomaceae species (n = 21) are documented from the Bahia state (Flora of Brazil, 2021). There have been few studies on Brazilian Cleomaceae species, and all were restricted to their tissue culture for pharmaceutical interests (Castro *et al.*, 2014; Simões, 2006) or regional floristic surveys (Akemi-borges and Pirani, 2017; Carneiro *et al.*, 2018). In addition, many of these species are endemic, and they may become extinct without being studied (*e.g.*, for *Haptocarpum*, there has been only one record corresponding to the type species, dated 1906, voucher B 100242671).

Studies on the floral morphology and reproductive biology of Brazilian Cleomaceae species are scarce, although such studies have been conducted in Africa (Zohoungbogbo *et al.*, 2018; Omondi *et al.*, 2017), North America (Cane, 2008; Higuera-Diaz *et al.,* 2015), and specific areas in Northeast Brazil (Machado et al., 2006). To this end, the present study furthers our understanding of the morphological diversity and reproductive biology of Cleomaceae species in different Brazilian biomes. The observed diversity of Cleomaceae would provide important insights into pollination and floral morphology within this family, mainly due to its phylogenetic proximity to Brassicaceae. Further integration of morphological, phylogenetic, ecological, and population genetic studies may shed light on the specific mechanisms underlying the major evolutionary transitions between reproductive modes (Barrett, 2002). In the present study, we also discuss the potential of the Cleomaceae family to serve as a research model for developing varieties and hybrids with improved floral traits, which would reduce dependence on pollinators or nectar and pollen production. Here, we describe diversity in the morphoanatomical characteristics of the flowers of 10 Cleomaceae species, in addition to respectively 7 and 5 accessions of *T. hassleriana* and *T. longicarpa*, and discuss how and to what extent the observed diversity contributes to the fitness of this plant family in relation to insect pollinators (or lack thereof). Specifically, we performed five types of pollination to verify the type that was the most efficient for seed set and explored the possible effects of these types on seed germination.

MATERIAL AND METHODS

Taxon sampling

The botanical material of Cleomaceae was collected from different Brazilian states (Fig. 1A, Table 1S). A list of the collected specimens and the details of their origins and herbarium vouchers are presented in Fig. 1A and Table 1S. The seeds of the 10 Cleomaceae species, 7 *T. hassleriana* accessions, and 5 *T. longicarpa* accessions were collected and germinated in plastic containers (volume, 5 L) with a commercial substrate. The plants (10 replicates per species) were grown in a greenhouse under semi-controlled conditions (maximum photosynthetically active radiation of 1,500 μmol photons m^{-2} s^{-1} and mean temperature of 30 ± 2°C) and daily irrigation. Under these conditions, the plants remained in the vegetative phase (3–5 months). Following the emergence of the flower bud, the plants were randomly placed in the botanical garden of the Universidade Federal de Viçosa (UFV) for approximately 5 months, until the pollination tests ended and the plants set fruits.

Floral morphology

The morphometric data of 50 mature and fully open flowers of all types were collected. Specifically, petal, sepal, stamen, ovary, gynophore, and androgynophore length was measured using a pachymeter and ruler. For sexing, 150 flowers were randomly collected on different days.

Reproductive biology

Fifty flowers per species and per pollination type were exposed to the treatments, totaling 5,000 tagged flowers. Reproductive success was evaluated based on the number of seeds produced through five different pollination types: (i) natural self-pollination (A): the buds of each flower were covered with paper bags to avoid insect access; (ii) open pollination (OP): insects were allowed to pollinate the flowers; (iii) hand self-pollination (HP): the flowers were pollinated with their own pollen; then the flowers were covered with paper bags to prevent cross-pollination; (iv) cross-pollination (CP): two flowers from two different plants were hand pollinated; the flowers have been emasculated; and (v) geitonogamy (G): pollen from the same plant was collected and applied to a different flower; the flowers have been emasculated.

Through the characterization of the flowers produced by 10 individuals of each species, we classified the type of sexual expression of each species according to Sakai and Weller (1999).

Additionally, we evaluated pollen limitation (PL), as described by Larson and Barrett, (2000), using the following formula: $PL = 1 - (OP/CP)$. A PL value exceeding 0.66 indicates high reproductive efficiency of the species resulting from the pollen flow promoted by pollinators (Zapata and Arroyo, 1978), while a value close to or below 0.25 indicates pollen limitation (Sobrevila and Arroyo, 1982). Values exceeding 1 are noted when the fruit set through open pollination is greater than that through cross-pollination, indicating highly efficient pollinators.

We also calculated the self-incompatibility index (ISI) using the following formula: $ISI = HP/CP$. The ISI is the ratio between the number of fruits formed through self-pollination and that through cross-pollination. As such, the ISI values below 0.25 indicates self-incompatibility (Bullock, 1985); values between 0.25 and 0.75 indicate self-incompatibility with some level of self-compatibility; and values exceeding 0.75 indicate self-compatibility (Lloyd and Schoen, 1992). Finally, we calculated the self-pollination index (AI) using the following formula: AI =

A/CP (Lloyd and Schoen, 1992). In 10 flowers that were previously bagged at the bud stage, nectar volume was measured as described by Machado *et al.,* (2006).

Seed germination test

Mature fruits (approximately 45 days after anthesis) produced from each tagged flower were harvested. Then, the seeds were removed from the fruit and allowed to dry at room temperature (~19 °C) for 5–7 days. Subsequently, the seeds were sown in a commercial substrate for the germination tests. Fifty seeds per species per treatment were used, which were separated into five pots with 10 seeds each. Finally, approximately 200–250 seeds per species were tested. The seeds were maintained in a growth chamber at the day/night temperature of 25°C/19°C under a 16:8 h photoperiod (~200 μmol photons $m^{-2} \cdot s^{-1}$) and 60% relative humidity. The germinated seed were counted weekly. As required, the substrate was moistened to maintain suitable conditions for germination.

Statistical analysis

The data were obtained from 15 flowers and/or fruits randomly tagged from 10 plants in individual pots per species, each representing a biological replicate. For germination tests, the data were obtained from five replicates, each with 10 seeds. The experiment followed a completely randomized design. The effect of the species was determined using analysis of variance ($P < 0.05$), and the means were subjected to Tukey’s test. Moreover, factor analysis and principal component analysis (PCA) were used to evaluate the associations among the different species. All statistical analyses were performed using Statistica (version 8.0) and R.

RESULTS

Diversity within Cleomaceae

The analyzed species presented diverse growth habits (herb or shrub), with compound leaves (3–7 leaflets), racemose inflorescences, and flowers of different colors (white or pale pink to purple) (Fig. 1A). Through daily observations, all species exhibited nocturnal anthesis and flower opening, beginning at sunset. At the beginning of anthesis, nectar—a floral reward offered—was secreted through a disk-shaped nectary. As such, large flowers (Spinosa I, Spinosa II, Cleoserrata, Parviflora, and Rosea) produced abundant nectar at the corolla base (~20–60 μL per flower) (Fig. 1S), and they could produce nectar multiple times (up to three times) within the same night. Meanwhile, small flowers (Aculeata and Gynandropsis) produced no more than 5 µL nectar per flower. In general, each inflorescence produced approximately five flowers daily, which could be of different sexes. Moreover, each inflorescence produced up to 900 µL of nectar (in large flowers) per day. Of note, a fully developed individual of the species branched profusely, bearing up to 10 racemes per plant.

These morphological characteristics, such as flower size, nectar production together with the sexual expression of the species and the production of pollen were used to conduct a principal component analysis (PCA) (Fig. 1B, Table 2S). The results showed the formation of four groups: (i) Spinosa + Parviflora + Rosea, (ii) Aculeata, (iii) Cleoserrata and (iv) Gynandropsis. The first two axes used for species separation explain 94.18 % of the variation. The contribution of each variable is shown in table 2S. Accordingly, the groupings were performed basically by the type of sexual expression of the species (next topic) (Fig. 1B)

Sexual expression in Cleomaceae species

Through the characterization of the flowers produced by 10 individuals of each species or access, it was possible to organize the 10 species in three groups of sexual expression: (i) andromonoecy, (ii) hermaphrodite and (iii) polygamous (Tables 1-3). Most of the species were allocated as polygamous species (Table 1), which means there are individuals able to produce all three types of flowers (hermaphrodite, male and female). However, it is worth mentioning that some polygamous species have unisexual female individuals and individuals with all three floral morphs (Table 2). The species categorized are *G. gynandra* (GG) - cluster Gynandropsis - and *T. hassleriana* accesses THC,

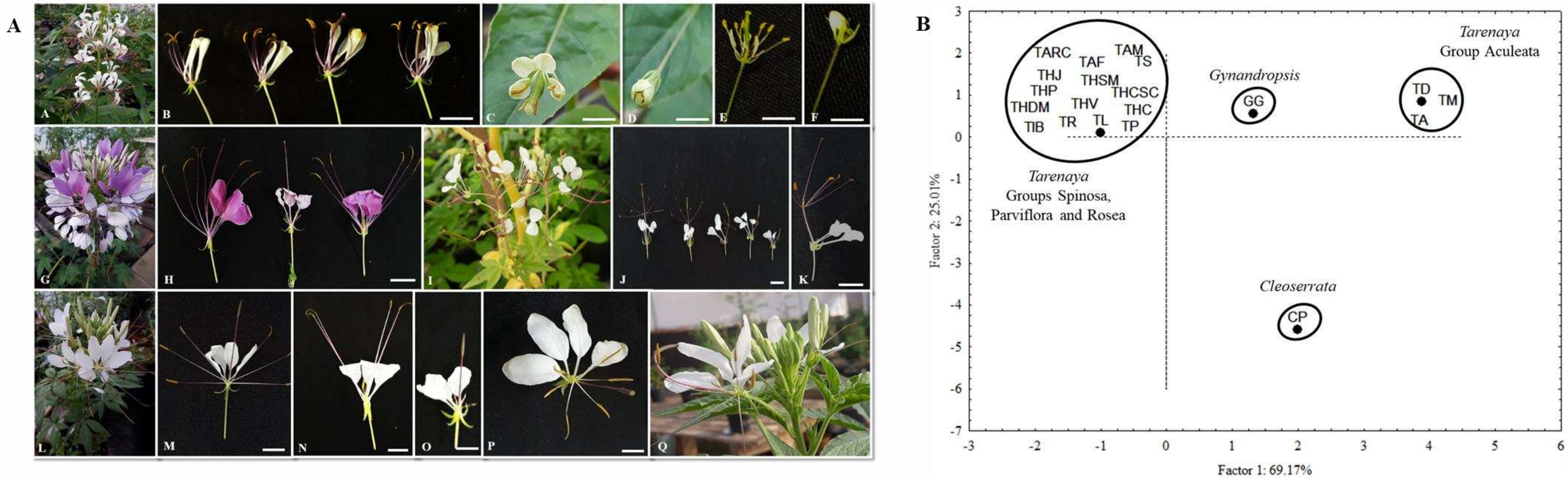
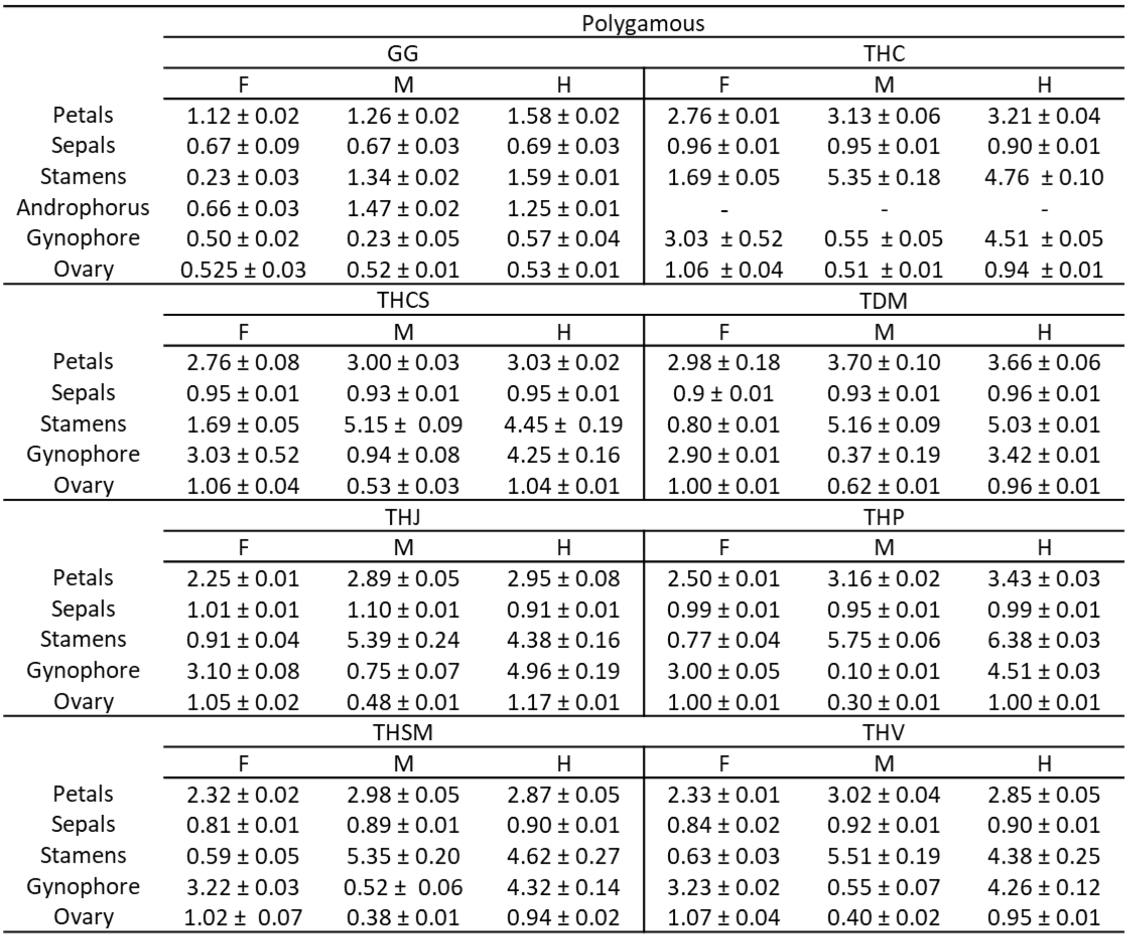


Fig. 1. Pictures of flowers and PCA based on floral features of Cleomaceae species. (**A**) - [Species name followed by Sexual Type] A-B *C. paludosa* (Andromonocy), A-inflorescence, B- staminate and hermaphrodite flowers. C-D *T. diffusa* (Hermaphrodite), C- Chasmogam, D-Cleistogam, E-F *T. microcarpa* (Hermaphrodite), E- Chasmogam, F-Cleistogam, G-H *T. hassleriana* (THDM) (Polygamous), G-inflorescence, H-B- hermaphrodite, pistillate and staminate flowers. I-K *G. gynandra* (Polygamous), I-inflorescence, J- hermaphrodite and pistillate, K- staminate flowers. L-O *T. spinosa* (Polygamous) L-inflorescence, M- hermaphrodite, N- staminate flowers, O- pistillate flowers. P-Q *T. longicarpa*, P- Hercogamy. O- Dicogamy. Bars=1cm. (**B**) **-** Principal component analysis based on floral characteristics such as sexual expression of the species (3), flower size and nectar production. Species (city and state of collection) - CP: *Cleoserrata paludosa* (Belém-PA); TA: *Tarenaya aculeata* (Feira de Santana-BA); TD: *T. diffusa* (Feira de Santana-BA); TM: *T. microcarpa* (Belém-PA).GG: *Gynandropsis gynandra* (Mossoró-RN); THC: *Tarenaya hassleriana* (Canaã-MG); THCS: *T. hassleriana* (Canoinhas-SC); THDM: *T. hassleriana* (Domingos Martins-ES); THJ: *T. hassleriana* (Joinville-SC); THP: *T. hassleriana* (Piau-MG); THS: *T. hassleriana* (São Miguel-MG); THV: *T. hassleriana* (Viçosa-MG); TL: *Tarenaya longicarpa* (Picos-PI); TR: *T. rosea* (Colatina-ES); TP: *T. parviflora* (Pombal-PB); TS: *T. spinosa* (Teresina-PI); TAF: *T. longicarpa* (Afrânio-PE); TAM: *T. longicarpa* (Manaus-AM); TARC: *T. longicarpa* (Arcoverde-PE); TIB: *T. longicarpa* (Ibimirim-PE).

THCSC, THDM, THJ, THP, THSM and THV, which belong to the cluster Spinosa II. In this group, female flowers are smaller than both male and hermaphrodite flowers, noticeable by length of the petals (Table 1). Stamens also have different lengh when compared amongst the sexual types of flowers. Occasionally, staminoides were observed in all the three sexual types, even though, they were always present in female flowers.

Table 1. Summary of petals, sepals, stamens, androphorus, gynophore and ovary lengths of the studied Cleomaceae species - Group polygamous. F: Pistillate flowers, M: Staminate flowers, H: Hermaphrodite flowers. Number corresponds to the medium value of the size (cm) of the structures, followed by ± SE, n=50.



	Polygamous					
	GG			THC		
	F	M	H	F	M	H
Petals	1.12 ± 0.02	1.26 ± 0.02	1.58 ± 0.02	2.76 ± 0.01	3.13 ± 0.06	3.21 ± 0.04
Sepals	0.67 ± 0.09	0.67 ± 0.03	0.69 ± 0.03	0.96 ± 0.01	0.95 ± 0.01	0.90 ± 0.01
Stamens	0.23 ± 0.03	1.34 ± 0.02	1.59 ± 0.01	1.69 ± 0.05	5.35 ± 0.18	4.76 ± 0.10
Androphorus	0.66 ± 0.03	1.47 ± 0.02	1.25 ± 0.01	-	-	-
Gynophore	0.50 ± 0.02	0.23 ± 0.05	0.57 ± 0.04	3.03 ± 0.52	0.55 ± 0.05	4.51 ± 0.05
Ovary	0.525 ± 0.03	0.52 ± 0.01	0.53 ± 0.01	1.06 ± 0.04	0.51 ± 0.01	0.94 ± 0.01
	THCS			TDM		
	F	M	H	F	M	H
Petals	2.76 ± 0.08	3.00 ± 0.03	3.03 ± 0.02	2.98 ± 0.18	3.70 ± 0.10	3.66 ± 0.06
Sepals	0.95 ± 0.01	0.93 ± 0.01	0.95 ± 0.01	0.9 ± 0.01	0.93 ± 0.01	0.96 ± 0.01
Stamens	1.69 ± 0.05	5.15 ± 0.09	4.45 ± 0.19	0.80 ± 0.01	5.16 ± 0.09	5.03 ± 0.01
Gynophore	3.03 ± 0.52	0.94 ± 0.08	4.25 ± 0.16	2.90 ± 0.01	0.37 ± 0.19	3.42 ± 0.01
Ovary	1.06 ± 0.04	0.53 ± 0.03	1.04 ± 0.01	1.00 ± 0.01	0.62 ± 0.01	0.96 ± 0.01
	THJ			THP		
	F	M	H	F	M	H
Petals	2.25 ± 0.01	2.89 ± 0.05	2.95 ± 0.08	2.50 ± 0.01	3.16 ± 0.02	3.43 ± 0.03
Sepals	1.01 ± 0.01	1.10 ± 0.01	0.91 ± 0.01	0.99 ± 0.01	0.95 ± 0.01	0.99 ± 0.01
Stamens	0.91 ± 0.04	5.39 ± 0.24	4.38 ± 0.16	0.77 ± 0.04	5.75 ± 0.06	6.38 ± 0.03
Gynophore	3.10 ± 0.08	0.75 ± 0.07	4.96 ± 0.19	3.00 ± 0.05	0.10 ± 0.01	4.51 ± 0.03
Ovary	1.05 ± 0.02	0.48 ± 0.01	1.17 ± 0.01	1.00 ± 0.01	0.30 ± 0.01	1.00 ± 0.01
	THSM			THV		
	F	M	H	F	M	H
Petals	2.32 ± 0.02	2.98 ± 0.05	2.87 ± 0.05	2.33 ± 0.01	3.02 ± 0.04	2.85 ± 0.05
Sepals	0.81 ± 0.01	0.89 ± 0.01	0.90 ± 0.01	0.84 ± 0.02	0.92 ± 0.01	0.90 ± 0.01
Stamens	0.59 ± 0.05	5.35 ± 0.20	4.62 ± 0.27	0.63 ± 0.03	5.51 ± 0.19	4.38 ± 0.25
Gynophore	3.22 ± 0.03	0.52 ± 0.06	4.32 ± 0.14	3.23 ± 0.02	0.55 ± 0.07	4.26 ± 0.12
Ovary	1.02 ± 0.07	0.38 ± 0.01	0.94 ± 0.02	1.07 ± 0.04	0.40 ± 0.02	0.95 ± 0.01

GG: *Gynandropsis gynandra* (Mossoró-RN); THC: *Tarenaya hassleriana* (Canaã-MG); THCS: *T. hassleriana* (Canoinhas-SC); THDM: *T. hassleriana* (Domingos Martins-ES); THJ: *T. hassleriana* (Joinville-SC); THP: *T. hassleriana* (Piau-MG); THS: *T. hassleriana* (São Miguel-MG); THV: *T. hassleriana* (Viçosa-MG).

The polygamous species *T. parviflora* (TP), *T. rosea* (TR), *T. longicarpa* (TL, TAF, TIB, TARC, and TAM), and *T. spinosa* (TS) were female or polygamous individuals (Table 2).

In this group, the female flowers of the female plants were smaller than those of the polygamous plants (Table 2). Meanwhile, the petals, sepals, stamens/staminodes, gynophores, and ovaries was shorter in the female plants. Of note, the polygamous, which comprised 16 species, differed in terms of the proportion of each flower type (male, hermaphrodite, and female) (Fig. 2S). Overall, these species bore more hermaphrodites and male flowers, although the occurrence of female flowers was evident (Fig. 2S).

Table 2. Summary statistics for mean values and standard deviations for petals, sepals, stamens, gynophore and ovary length in accesses of Cleomaceae species – Group Polygamous (unisexual female individual and individual with the three floral morphs - female, male and hermaphrodite). F: Pistillate flowers, M: Staminate flowers, H: Hermaphrodite flowers. Number corresponds to the medium value of the size (cm) of the structures, followed by ± SE, n=50.

	Polygamous					
	GG			THC		
	F	M	H	F	M	H
Petals	1.12 ± 0.02	1.26 ± 0.02	1.58 ± 0.02	2.76 ± 0.01	3.13 ± 0.06	3.21 ± 0.04
Sepals	0.67 ± 0.09	0.67 ± 0.03	0.69 ± 0.03	0.96 ± 0.01	0.95 ± 0.01	0.90 ± 0.01
Stamens	0.23 ± 0.03	1.34 ± 0.02	1.59 ± 0.01	1.69 ± 0.05	5.35 ± 0.18	4.76 ± 0.10
Androphorus	0.66 ± 0.03	1.47 ± 0.02	1.25 ± 0.01	-	-	-
Gynophore	0.50 ± 0.02	0.23 ± 0.05	0.57 ± 0.04	3.03 ± 0.52	0.55 ± 0.05	4.51 ± 0.05
Ovary	0.525 ± 0.03	0.52 ± 0.01	0.53 ± 0.01	1.06 ± 0.04	0.51 ± 0.01	0.94 ± 0.01
	THCS			TDM		
	F	M	H	F	M	H
Petals	2.76 ± 0.08	3.00 ± 0.03	3.03 ± 0.02	2.98 ± 0.18	3.70 ± 0.10	3.66 ± 0.06
Sepals	0.95 ± 0.01	0.93 ± 0.01	0.95 ± 0.01	0.9 ± 0.01	0.93 ± 0.01	0.96 ± 0.01
Stamens	1.69 ± 0.05	5.15 ± 0.09	4.45 ± 0.19	0.80 ± 0.01	5.16 ± 0.09	5.03 ± 0.01
Gynophore	3.03 ± 0.52	0.94 ± 0.08	4.25 ± 0.16	2.90 ± 0.01	0.37 ± 0.19	3.42 ± 0.01
Ovary	1.06 ± 0.04	0.53 ± 0.03	1.04 ± 0.01	1.00 ± 0.01	0.62 ± 0.01	0.96 ± 0.01
	THJ			THP		
	F	M	H	F	M	H
Petals	2.25 ± 0.01	2.89 ± 0.05	2.95 ± 0.08	2.50 ± 0.01	3.16 ± 0.02	3.43 ± 0.03
Sepals	1.01 ± 0.01	1.10 ± 0.01	0.91 ± 0.01	0.99 ± 0.01	0.95 ± 0.01	0.99 ± 0.01
Stamens	0.91 ± 0.04	5.39 ± 0.24	4.38 ± 0.16	0.77 ± 0.04	5.75 ± 0.06	6.38 ± 0.03
Gynophore	3.10 ± 0.08	0.75 ± 0.07	4.96 ± 0.19	3.00 ± 0.05	0.10 ± 0.01	4.51 ± 0.03
Ovary	1.05 ± 0.02	0.48 ± 0.01	1.17 ± 0.01	1.00 ± 0.01	0.30 ± 0.01	1.00 ± 0.01
	THSM			THV		
	F	M	H	F	M	H
Petals	2.32 ± 0.02	2.98 ± 0.05	2.87 ± 0.05	2.33 ± 0.01	3.02 ± 0.04	2.85 ± 0.05
Sepals	0.81 ± 0.01	0.89 ± 0.01	0.90 ± 0.01	0.84 ± 0.02	0.92 ± 0.01	0.90 ± 0.01
Stamens	0.59 ± 0.05	5.35 ± 0.20	4.62 ± 0.27	0.63 ± 0.03	5.51 ± 0.19	4.38 ± 0.25
Gynophore	3.22 ± 0.03	0.52 ± 0.06	4.32 ± 0.14	3.23 ± 0.02	0.55 ± 0.07	4.26 ± 0.12
Ovary	1.02 ± 0.07	0.38 ± 0.01	0.94 ± 0.02	1.07 ± 0.04	0.40 ± 0.02	0.95 ± 0.01

TL: *Tarenaya longicarpa* (Picos-PI); TR: *T. rosea* (Colatina-ES); TP: *T. parviflora* (Pombal-PB); TS: *T. spinosa* (Teresina-PI); TAF: *T. longicarpa* (Afrânio-PE); TAM: *T. longicarpa* (Manaus-AM); TARC: *T. longicarpa* (Arcoverde-PE); TIB: *T. longicarpa* (Ibimirim-PE).

Table 3. Summary statistics for mean values and standard deviations for petals, sepals, stamens, gynophore, ovary and floral bud length in accesses of Cleomaceae species – Group Andromonoecy and Hermaphrodite. M: Staminate flowers, H: Hemaphrodite flowers. Cleis: cleistogamous, Casm: chasmogamous flowers. Number corresponds to the medium value of the size (cm) of the structures, followed by ± SE, n=50.

	Andromonoecy		Hermaphrodite					
	CP		TA		TD		TM	
	M	H	Cleis	Casm	Cleis	Casm	Cleis	Casm
Petals	2.00 ± 0.01	1.96 ± 0.01	-	1.25 ± 0.01	-	1.20 ± 0.01	-	1.23 ± 0.01
Sepals	0.41 ± 0.01	0.50 ± 0.01	-	0.40 ± 0.01	-	0.40 ± 0.01	-	0.41 ± 0.01
Stamens	1.93 ± 0.01	1.95 ± 0.01	-	0.69 ± 0.01	-	0.67 ± 0.01	-	0.66 ± 0.01
Gynophore	0.1 ± 0.01	0.62 ± 0.01	-	0.35 ± 0.01	-	0.32 ± 0.01	-	0.33 ± 0.01
Ovary	0.4 ± 0.01	0.84 ± 0.01	-	0.45 ± 0.01	-	0.44 ± 0.01	-	0.44 ± 0.01
Floral bud	0.00	0.00	0.52 ± 0.01	0	0.41 ± 0.01	-	0.47 ± 0.01	-

CP: *Cleoserrata paludosa* (Belém-PA); TA: *Tarenaya aculeata* (Feira de Santana-BA); TD: *T. diffusa* (Feira de Santana-BA); TM: *T. microcarpa* (Belém-PA).

C. paludosa alone was classified as andromonoecious, which comprised accessions that produced hermaphrodite and male flowers on the same plant (Table 3). The flowers were small, without evident differences between the male and hermaphrodite flowers, except in the length of the gynophores and ovaries (Table 3).

Finally, the hermaphrodite species *T. aculeata* (TA), *T. microcarpa* (TM), and *T. diffusa* (TD) produced only hermaphrodite flowers. Nevertheless, these flowers could be chasmogamous (open flower for cross-pollination) or cleistogamous (non-opening flowers for self-pollination) (Fig. 3S, Table 3).

Effects of pollination type on fruit set

Plants may employ specific strategies to avoid autogamy; however, self-compatible plants may follow geitonogamy and self-pollination. In the present study, we verified two strategies of plants to avoid self-pollination, namely dichogamy and hercogamy. In dichogamy, the androecium and gynoecium mature at different times, with both types observed in all species, except in the Aculeata cluster. In some of the observations, we found that the gynoecium matured up to 12 hours after the androecium (the gynoecium was covered by a petal, preventing the crossing). However, most of the time the difference in maturation between the organs was up to an hour. In addition, most of the Cleomaceae species exhibited hercogamy (Fig. 1A). Notably, the studied species are self-compatible (ISI > 0.3) (Fig. 4S); therefore, various types of mating systems may occur given the lack of pollen limitation (PL > 0.20).

However, only GG, TA, TD, and TM exhibited autogamy (AI > 0.7), while the rest of the species exhibited AI values below 0.7 (Fig. 4S), indicating the lack of self-pollination.

Overall, hand cross-pollination produced the highest seed yield, followed by open pollination (Figs. 2-3). Conversely, natural self-pollination (autogamy) produced the lowest seed yield (n = 18), with no seed production in 16 species (Figs. 2-3). Through natural self-pollination, TD produced the highest number of seeds. Through cross-pollination, TA, THC, THV, THP, THCSC, THSM, TL, TR, TS, TAF, TAM, and TIB produced the highest number of seeds (Figs. 2-3). GG was the only species that produced the highest number of seeds through open pollination. In contrast, THDM and TP produced more seeds per silique through geitonogamy. *C. paludosa*, TM, and THJ produced more seeds through cross- and open pollination, although there were no significant differences between these treatments. TARC produced more seeds through cross-pollination, open pollination, and geitonogamy (Figs. 2-3).

Seeds germination

Species in the Spinosa I cluster exhibited the highest germination rate (~65%), followed by those in Gynandropsis and Cleoserrata (both 38%), Spinosa II (35%), and Aculeata (25%) (Fig. 5S). In addition, the seeds of Spinosa I species germinated faster and were more homogeneous than those of the other species. However, there were no significant differences in the mean germination rate among the pollination treatments: CP = 44%, G = 41%, HP = 40%, and OP = 55% (Fig. 5S). Of note, with the exception of the herbaceous species GG, TA, TD, and TM, which exhibited 32%, 24%, 18%, and 22% germination following autogamy, no other species set fruit under this treatment (Figs. 4-5).

Overall, for most species, there were no significant differences in germination rate among the different treatments (Figs. 4-5). Specifically, there were no significant differences in the seed germination of TA, TD, and TM (~25%) among the five pollination treatments. Likewise, there were no significant differences in the seed germination of TARC (57%), TR (65%), THCS (37%), THS (32%), and TAM (75%) among the five pollination treatments, except in autogamy, under which no seeds were produced (Figs. 4-5).

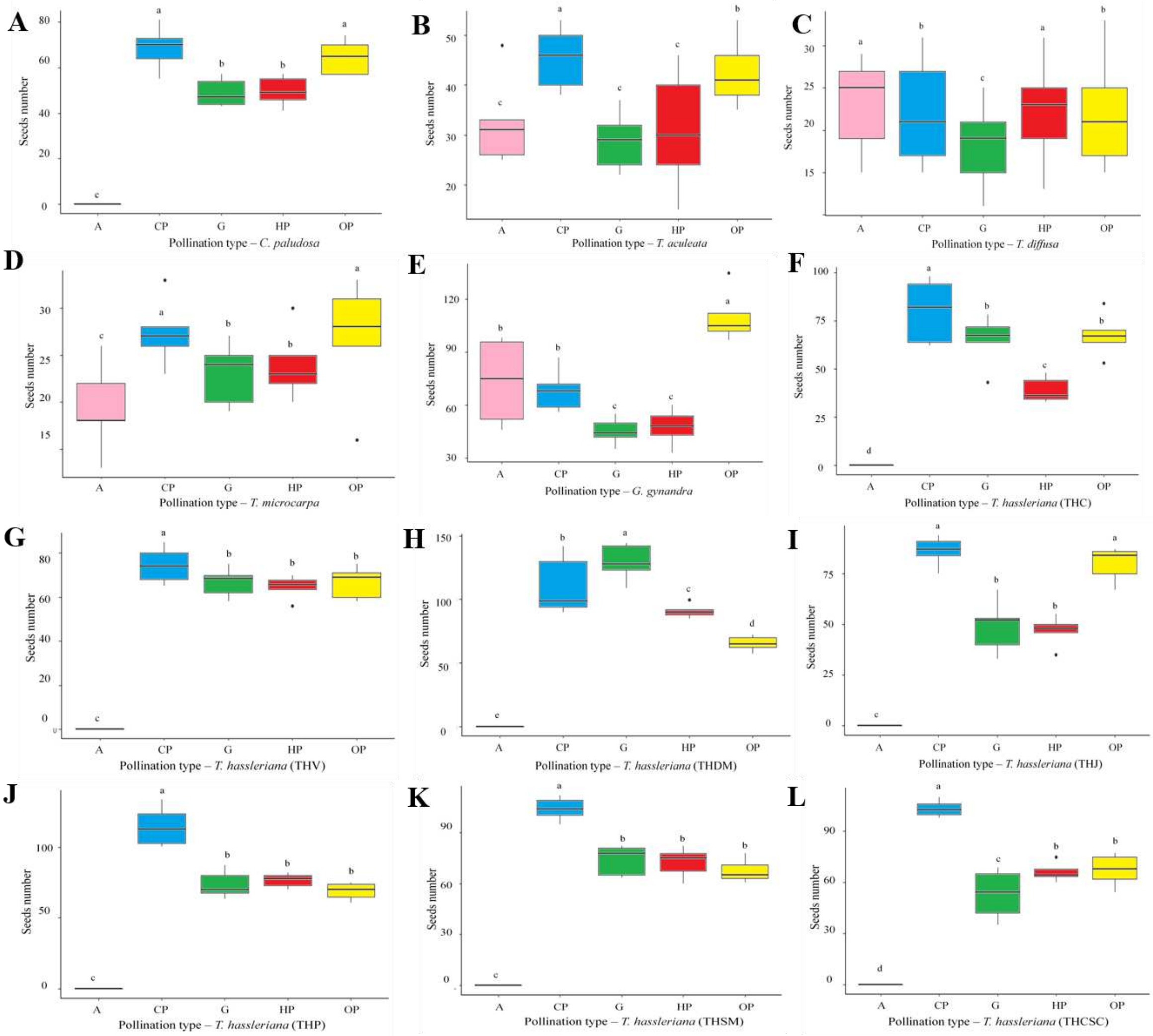


Fig. 2. Mature seeds per fruit according to the pollination treatments. Letters above individual box-scatter indicate significant groupings according to Tukey's Test, n=50. The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. Natural self-pollination (A); Open pollination (OP); Hand self-pollination (HP); Cross-pollination (CP); Geitonogamy (G).

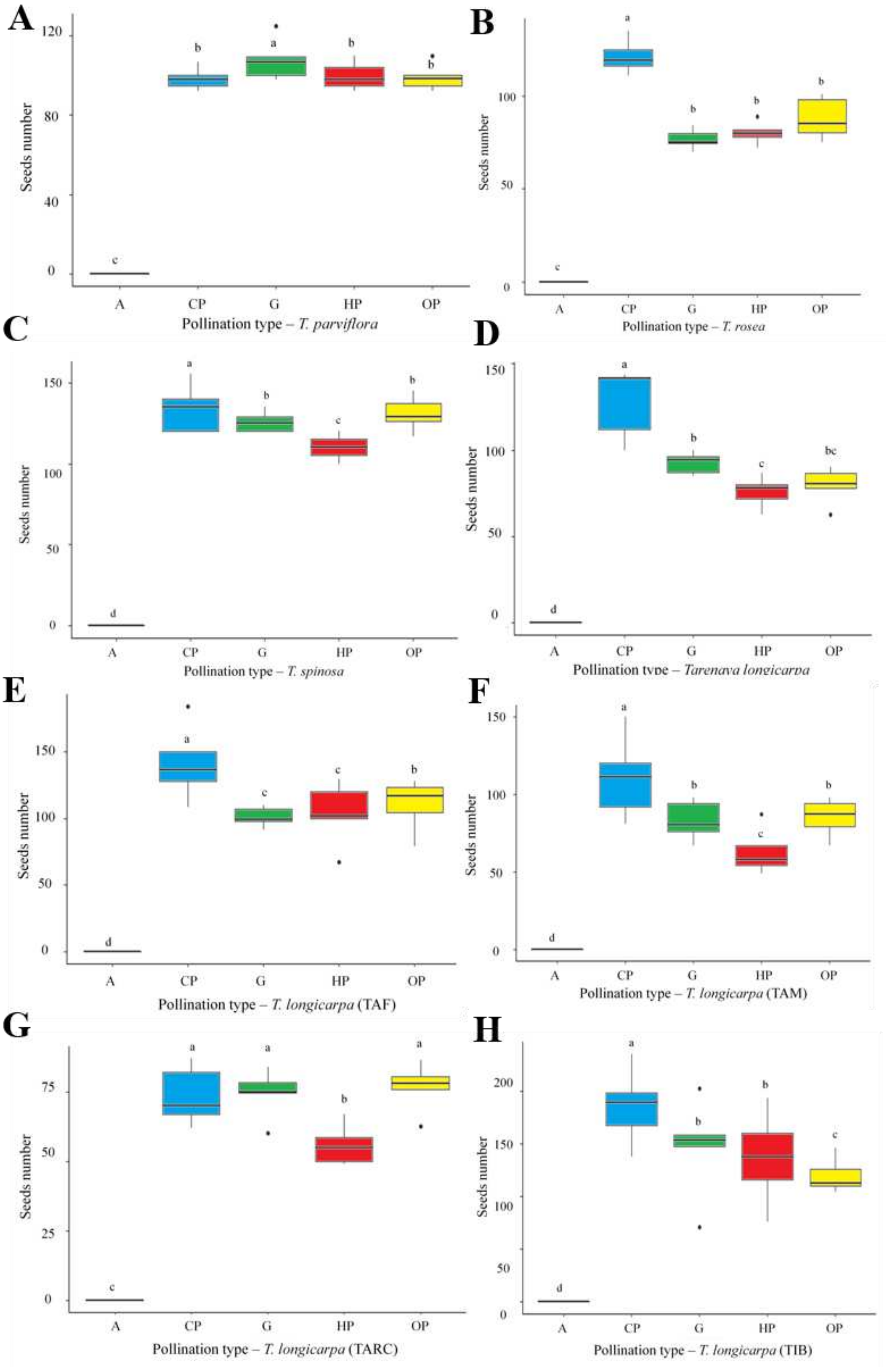


Fig. 3. Mature seeds per fruit by pollination treatments for Cleomaceae species. Letters above individual box-scatter indicate significant groupings according to Tukey's Test, n=50. The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. Natural self-pollination (A); Open pollination (OP); Hand self-pollination (HP); Cross-pollination (CP); Geitonogamy (G).

DISCUSSION

Cleomaceae exhibits remarkable diversity in floral morphology and reward production

Hercogamy reduces the interference between the sexual parts and lowers the deposition of self-pollen on the stigma, thereby minimizing the conflict between sexes in hermaphrodite flowers (Barrett, 2002). Heterostyly is one of the most studied characteristics of the Rubiaceae family, which is characterized by a higher number of distylous species than any other botanical family (Matias *et al.*, 2016). In contrast, although Cleomaceae includes many heterostylous species, unlike Rubiaceae, which includes distylous species (Matias *et al.*, 2016), Cleomaceae species exhibit several floral morphs. In the present study, we observed both functionally staminate short-gynoecium floral types and functionally hermaphrodite medium and long-gynoecium floral types (Fig. 1), consistent with the findings in *G. gynandra* reported by Zohoungbogbo et al. (2018). Of note, these characteristics may be specifically and directly linked to habitat loss and the different selection pressures imposed on isolated forest fragments (Barrett, 2002).

We observed heterostyly in the studied Cleomaceae species (Fig. 1A), except for those in the Aculeata cluster, as previously reported in *G. gynandra* (Zohoungbogbo *et al.*, 2018). This trait has likely evolved multiple times in Cleomaceae, since some species do not present heterostyly (*e.g.*, *C. brachycarpa*, *C. violacea*, and *C. karachiensis*, which present homostyly; Riaz et al., 2019). In the present study, the heterostylous species exhibited a high degree of self-compatibility (ISI > 0.30) (Fig. 2S), which may be related to reproductive safety, particularly when populations are under certain selective pressure (Schoen *et al.*, 1997). Thus, in addition to studies on the reproductive biology and floral morphometrics, the frequency of specific floral morphs in the populations of heterostylous species must be explored to understand the adaptive meaning of this polymorphism (Matias *et al.*, 2016). In the present study, for instance, *G. gynandra* possessed 27% female, 10% male, and 63% of hermaphrodite flowers (Fig. 1S), contrary to the observations in the study by Zohoungbogbo et al., (2018) in the same species but from other population (70% male and 30% hermaphrodite flowers).

The factors conferring the observed plasticity in sex expression in Cleomaceae remain unknown; nonetheless, in addition to the endogenous factors, such as hormones (*e.g.,* autonomous gibberellin pathways), environmental factors, such as photoperiod, vernalization, temperature, drought, and salinity, exogenous applied hormones and chemicals, and pathogenic microbes influence sex expression in many species (Banks, 2008; Chuck et al., 2007; Lai et al., 2018; Pawełkowicz et al., 2019; Aryal et al., 2014; She et al., 2009). In this sense, *Cucumis sativus* is the most studied species (Lai et al., 2018; Pawełkowicz et al., 2019), although there has been some research on other species, such as *Silene latifolia* (Muyle et al., 2012), *Asparagus officinalis* (Telgmann-Rauber *et al.*, 2007; Wang *et al.*, 2014), *Fragaria elatior* (Vyskot and Hobza, 2015, 2004), *Melandrium album* (Janousek *et al.,* 1996), and *Elaeis guineensis* (Jaligot et al., 2011), as the primary signals inducing sex development and the signal transduction pathways differ between across (for review, see Pawełkowicz *et al.,* 2019). Therefore, given the variability in factors that regulate and influence floral formation, the elucidation of the molecular mechanisms underlying sex determination represents an enormous challenge in plant biology (Pawełkowicz *et al.,* 2019). Of note, environmental factors affect the sex of flowers as well as the timing of flowering (Kozlov *et al.*, 2020; Cho *et al.*, 2017). Therefore, mechanisms underlying flower formation warrant further exploration to aid the development of novel techniques of sex manipulation, particularly for crops. Previous studies in this regard were very broad, analyzing the genomics and transcriptomics of flower buds and flowers and exploring the links among the related genes (Durand and Durand, 1991).

Cleomaceae employs various pollination strategies

Cleomaceae species benefit from both self- and cross-pollination to ensure reproductive success. However, hand cross-pollination and open pollination resulted in the highest seed yield in most of the Cleomaceae species studied (Figs. 2 and 3). This finding reflects the importance of pollinators in plant reproductive mechanisms, enabling gene flow between individuals and increasing genetic variability. Cross-pollination improved crop quality and yield, in addition to enhancing seed and fruit set (Garibaldi *et al.*, 2009; Potts *et al.*, 2010; Vinícius-Silva *et al.*, 2017; Elisante *et al.,* 2020; Chen and Zhao 2019). Meanwhile, as observed in the Aculeata and Gynandropsis clusters (Fig. 2), regardless of the genetic benefits of outcrossing, a significant number of angiosperm species (around 20%) have evolved to be predominantly autogamous (Barrett, 2002). However, this trait may be highlighted as an advantage in adverse environments with low frequency of floral visitors, as anticipated in the face of current climate crisis. Therefore, Cleomaceae is a key model to understand the various pollination strategies, given that the populations of pollinators are currently dwindling (Hallmann *et al.*, 2017; Marshman *et al.*, 2019) and at least 75% of plants still dependent on pollinators (Klein *et al.*, 2007; Kremen, 2018).

Furthermore, the type of pollination did not affect seed germination (Figs. 4-5). Interestingly, however, the recorded germination rates (of the untreated seeds) of majority of the species were different from the values reported previously. For instance, for *G. gynandra*, we recorded the germination rate of 38%, whereas Zohoungbogbo et al. (2018) and Ekpong (2009) have recorded the germination rate of only 25% and 17%, respectively. Overall, expect in the Spinosa I cluster, the germination rate remained low in Gynandropsis, Cleoserrata, Spinosa II, and Aculeata, which may be explained by the fact that the mature seeds of Cleomaceae species enter a period of inactivity or extremely low metabolic activity (dormancy), as already verified in *G. gynandra* (Zohoungbogbo *et al.*, 2018; Ekpong, 2009), *T. spinosa* (Castro *et al.*, 2014), *T. hassleriana*, *Polanisia dodecandra* (Gomez Raboteaux and Anderson, 2010), *Cleome lutea*, and *C. serrulata* (Cane, 2008). Nonetheless, the higher germination rates observed in these previous studies were achieved through additional treatments, such as prolonged storage at specific temperatures; application of GA3 or KNO_3; and leaching, pre-chilling, soaking, and pre-heating at specific temperatures (e.g., vernalization). In addition, the seeds of some Cleomaceae species (e.g., *T. hassleriana*) bear thick coats, which delay germination. In such species, scarification, which involves mechanical or chemical processes to soften the seed coat, can improve the germination rate.

Moreover, although no pollen incompatibility was detected in the crosses of the studied species, the overall low germination rate of Cleomaceae species may be attributed to the disruption of the pollen, embryo sac, embryo, and endosperm development, which may result in seed abortion (Shao *et al.*, 2020). Normal viable seeds cannot form during pollination, fertilization, and embryo development (Shao *et al.*, 2020). Thus, a systematic study of these processes may provide a theoretical basis and reference for exploring the mechanism of seed abortion and establishing an efficient assisted-breeding system.

Perspectives on Cleomaceae

The species studied here reflect the remarkable diversity of the Cleomaceae family in terms of morphology (Patchell *et al.*, 2011), photosynthetic mechanisms (Feodorova et al., 2010; Marshall et al., 2007; Voznesenskaya et al., 2007), genomics (Inda *et al.*, 2008), pollination (Cane, 2008; Machado et al., 2006; Zohoungbogbo et al., 2018), and geographic distribution (Bayat *et al.*, 2018). Accordingly, this family has been proposed as a model thanks to its several peculiarities and its close phylogenetic relationship with *Arabidopsis thaliana* of the Brassicaceae family, from which Cleomaceae has diverged relatively recently (~35 mya.) (Schranz and Mitchell-olds, 2006).

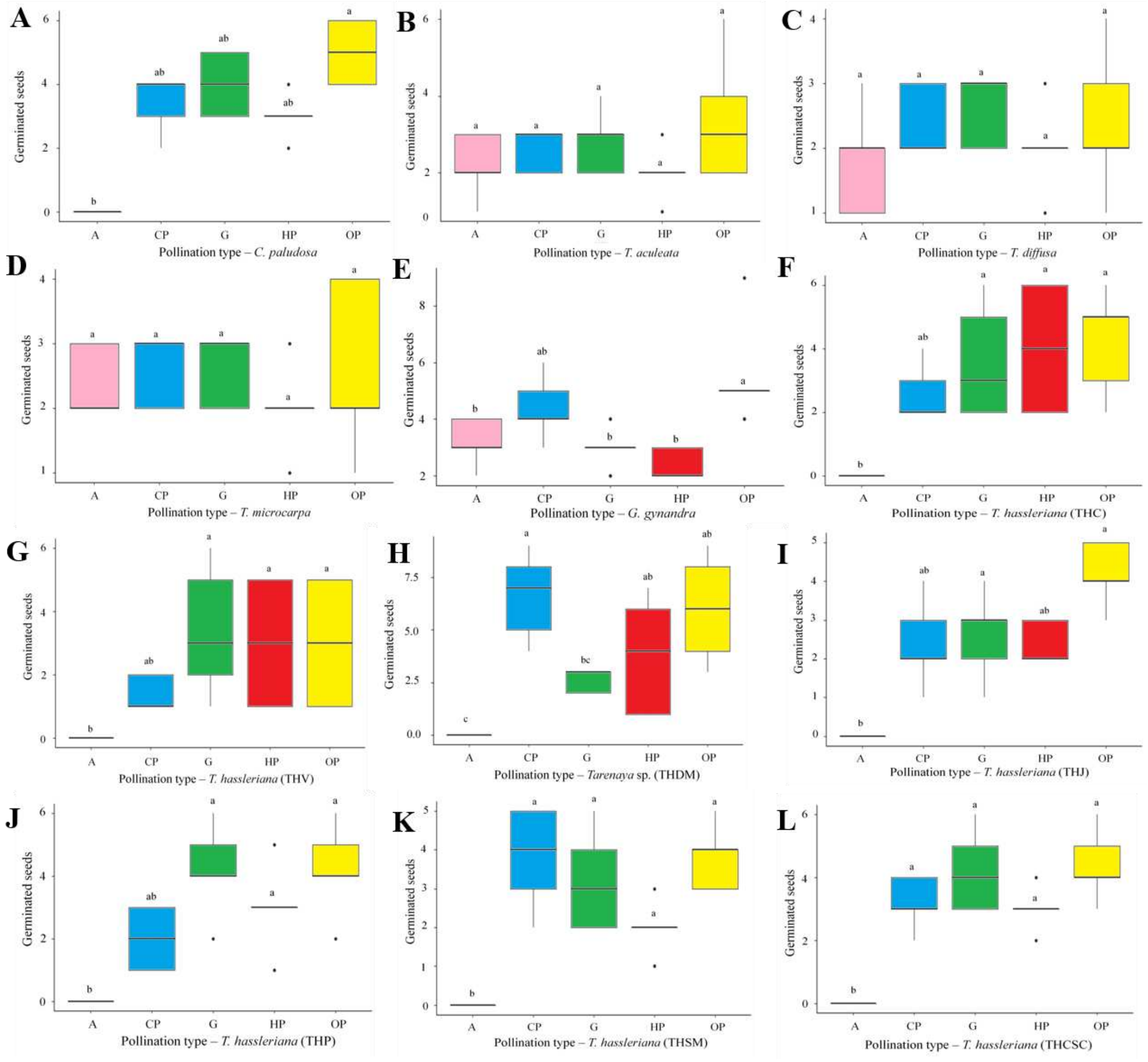


Fig. 4. Germinated seeds by pollination treatments for Cleomaceae species. Letters above individual box-scatter indicate significant groupings according to Tukey's Test, n=5. Solid lines in each box indicate the median; data dispersion is represented by the interquartile range, followed by standard error and outliers. Natural self-pollination (A); Open pollination (OP); Hand self-pollination (HP); Cross-pollination (CP); Geitonogamy (G).

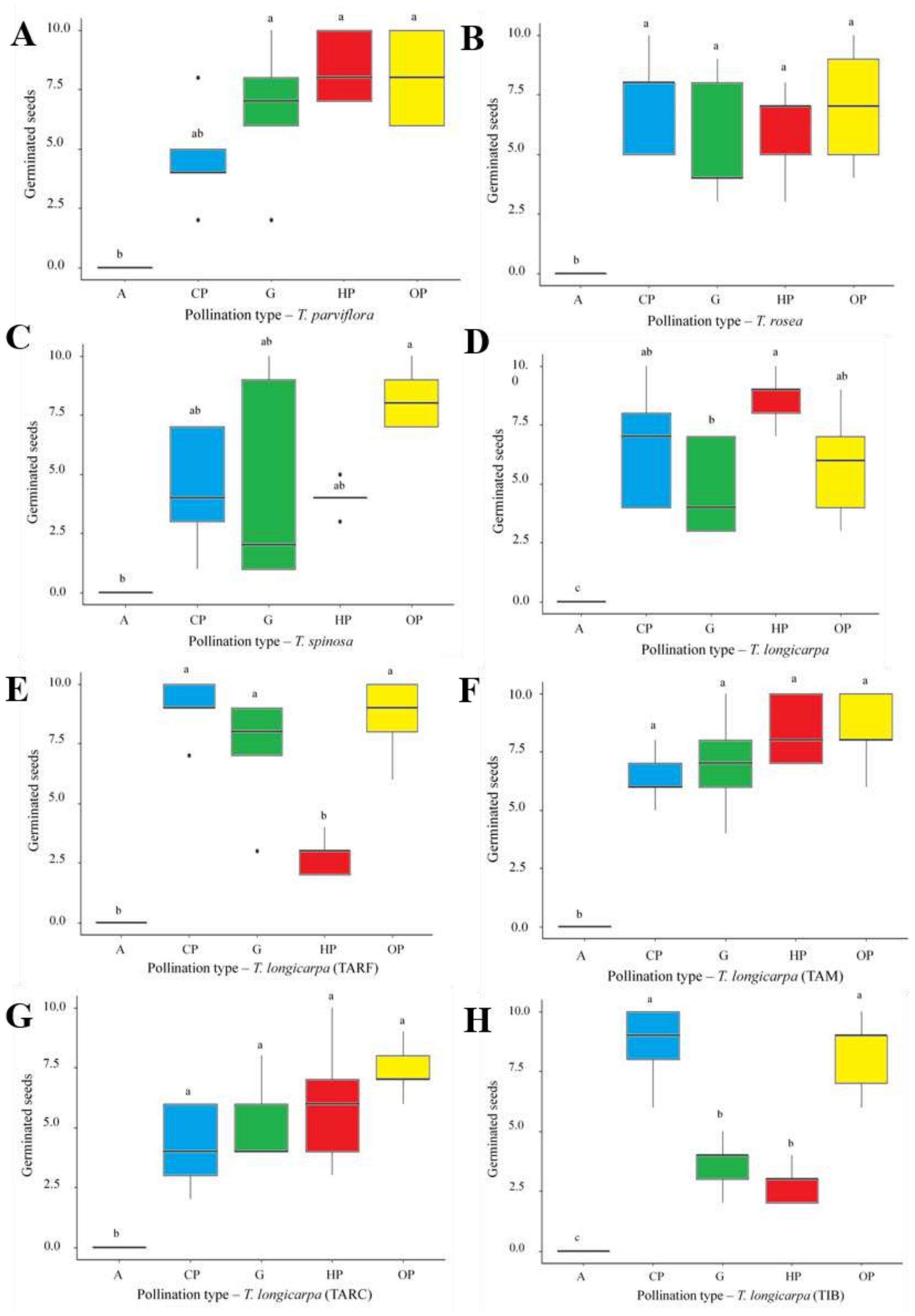


Fig. 5. Germinated seeds by pollination treatments for Cleomaceae species. Letters above individual box-scatter indicate significant groupings according to Tukey's Test, n=5. The median is indicated by solid lines in each box, data dispersion is represented by the interquartile range, followed by standard error and outliers. Natural self-pollination (A); Open pollination (OP); Hand self-pollination (HP); Cross-pollination (CP); Geitonogamy (G).

MADS-box genes are the key regulators of several aspect of plant reproductive development, serving pivotal functions in the regulation of flowering time, inflorescence architecture, floral organ identity, and seed development (Causier *et al.*, 2010; Lai *et al.*, 2021; Schilling *et al.*, 2018). Furthermore, according to the (A)BCE model, these genes synergistically control the identity of four different floral organs (Causier *et al.*, 2010). The paralogs of A, B, C, and E have been explored in several species to elucidate, in addition to the function, their level of divergence among plant families (Bradley *et al.*, 1993; Davies *et al.*, 1999; Liljegren *et al.*, 2000; Pelaz *et al.*, 2000; Jack *et al.*, 1992; Bruijn *et al.*, 2013; de Bruijn *et al.*, 2018). Specifically, *T. hassleriana* (Cleomaceae) harbors two copies of the floral B-class gene *PI*, one of which is similar to that found in other angiosperms and the other is related to that present in the Brassicaceae members (de Bruijn *et al.*, 2018). In addition to differences among families, the copy number and expression patterns of genes vary even within the same family. For instance, *T. hassleriana* harbors two highly similar *AP3* paralogs originating from a recent tandem duplication, which is lacking in *G. gynandra* (de Bruijn *et al.*, 2018).

Therefore, further studies on the MADS-box gene family may allow us to discover the transcription factors responsible for the diversity of floral morphology and sex expression in Cleomaceae. We have a long way to go yet to understand the variability observed in the family. Although there is some evidence of the supergene control of the morphological and physiological components of heterostyly, virtually nothing is known regarding this (Barrett, 2002). Clarifying the factors governing floral diversity will allow us to induce (e.g., through the induction of hercogamy or production of unisexual flowers) or even avoid cross-pollination (e.g., through the induction of self-pollination).

Floral rewards, such as nectar, pollen, and resin, facilitate crop pollination, thereby increasing the amount and breadth of food available to humans (Prasifka *et al.,* 2018). Most of the species here studied, in addition to bats, are pollinated by bees (e.g., Machado *et al.*, 2006), which benefit both from the large amounts of pollen (proteins) and nectar (sugars) produced. The abundance and diversity of pollinators, particularly bees, have declined over the past several decades (Aizen *et al.*, 2009; Hallmann *et al.*, 2017), eliciting cascading effects on the food webs and jeopardizing the ecosystem services (Hallmann *et al.*, 2017). In this regard, the creation of flower-rich habitats, such as hedgerows, field borders, or cover crops, has been proposed to conserve bee populations and enhance crop pollination (Williams *et al.*, 2015; Wratten *et al.*, 2012). Nonetheless, whether these habitats actually increase the number of pollinators required for targeted crop pollination (Sidhu and Joshi, 2016) and whether flowers in these habitats compete with the crop to interfere with crop pollination (Bostanian *et al.,* 2013) remain unclear. Hence, to help address the needs of agriculture and pollinator preservation, the development of crop varieties with specific nectar or nectar-related traits (as observed in Cleomaceae species) for attracting and retaining pollinating insects is an appealing strategy (Prasifka *et al.,* 2018). Thus, Cleomaceae (which is phylogenetically close to Brassicaceae) has emerged as a promising model to devise strategies for improving crop yield (through promoting cross-pollination) and floral reward production (pollen and nectar) for pollinators, particularly bees.

CONCLUSION

Cleomaceae species have different types of sexual expression, essentially characterized by three types: andromonoecy, hermaphrodite plants and polygamous. In addition, through the various pollination tests, the perception of a need for pollinators was confirmed. The majority of species had the higher yield in number of seeds when cross pollinated. Evolutionary features, such as hercogamy, dichogamy and unisexual flowers, to avoid self-pollinated were present in some of them. All of that lead us to a co-evolution perspective through the Cleomaceae family. Furthermore, it should be noted that the floral diversity and production of rewards observed are important characteristics at times of pollinators population decline, since the larger part of angiosperms is pollinated by animals, especially bees. In this way, the production of food is also linked to pollinators once they improve food quality and quantity. In this sense, Cleomaceae family expressed interesting tools to attract pollinators, what could be very useful if inserted in crop species by isolating interest genes. Further studies could check the viability of those factors regarding plant breeding programs.

ACKNOWLEDGMENTS

We thank Prof. Dr. Jocelyn Hall for her insightful critique of and contributions to this manuscript. This work was supported by funding from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; grant number 424024/2018-7) and the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG; grant numbers APQ-00528-18 and CRA-RED00053-16). ANN and WLA acknowlegde the research fellowships granted by CNPq. DFP and KFS acknowlegde the scholarship granted by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). MGMVV acknowlegdes the scholarship granted by CAPES (PNPD-1638006). APMW acknowledges funding from the

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy (EXC-2048/1; project ID 390686111).

LITERATURE CITED

Aizen, M.A., Garibaldi, L.A., Cunningham, S.A. and Klein, A.M. (2009) How much does agriculture depend on pollinators? Lessons from long-term trends in crop production. *Ann. Bot.*, **103**, 1579–1588.

Akemi-borges, I. and Pirani, J.R. (2017) Flora of Serra do Cipó, Minas Gerais: Cleomaceae. *Boletim de Botânica*, 0–6.

Aryal R., Jagadeeswaran G., Zheng Y, Yu Q., Sunkar R., Ming R. (2014) Sex specific expression and distribution of small RNAs in papaya. *BMC Genomics*, **15**.

B Janousek 1, J Siroký, B Vyskot (1996) Epigenetic control of sexual phenotype in a dioecious plant, Melandrium album. *Mol. Gen. Genet.*, **250**, 483–490.

Boris Vyskot, Roman Hobza (2004) Gender in plants: sex chromosomes are emerging from the fog. *Trends Genet.*, **20**, 432–438.

Boris Vyskot, Roman Hobza (2015) The genomics of plant sex chromosomes. *Plant Sci.*, **236**, 126–135.

Bailes, E.J., Ollerton, J., Pattrick, J.G. and Glover, B.J. (2015) How can an understanding of plant-pollinator interactions contribute to global food security? *Curr. Opin. Plant Biol.*, **26**, 72–79.

Banks, J.A. (2008) MicroRNA, sex determination and floral meristem determinacy in maize. *Genome Biol.*, **9**, 204.

Barrett, S.C.H. (2002) The evolution of plant sexual diversity. *Nat. Rev. Genet.*, **3**, 274–284.

Bayat, S., Schranz, M.E., Roalson, E.H. and Hall, J.C. (2018) Lessons from Cleomaceae , the Sister of Crucifers. *Trends Plant Sci.*, 1–14.

Biesmeijer, J.C., Roberts, S.P.M., Reemer, M., et al. (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science (80)*, **313**, 351–354.

Bradley, D., Carpenter, R., Sommer, H., Hartley, N. and Coen, E. (1993) Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the plena locus of antirrhinum. *Cell*, **72**, 85–95.

Brittain, C.A., Vighi, M., Bommarco, R., Settele, J. and Potts, S.G. (2010) Impacts of a pesticide on pollinator species richness at different spatial scales. *Basic Appl. Ecol.*, **11**, 106–115.

Brown, N.J., Parsley, K. and Hibberd, J.M. (2005) The future of C4 research – maize, Flaveria or Cleome? *Trends Plant Sci.*, **10**, 215–221.

Bruijn, S. De, Bhide, A.S., Kuelahoglu, C., et al. (2013) The *Tarenaya hassleriana* Genome Provides Insight into Reproductive Trait and Genome Evolution of Crucifers. *The Plant Cell*, **25**, 2813–2830.

Bruijn, S. de, Zhao, T., Muiño, J.M., Schranz, E.M., Angenent, G.C. and Kaufmann, K. (2018) PISTILLATA paralogs in Tarenaya hassleriana have diverged in interaction specificity. *BMC Plant Biol.*, **18**, 1–18.

Bullock, S.H. (1985) Breeding Systems in the Flora of a Tropical Deciduous Forest in Mexico. *Biotropica*, **17**, 287.

Burkle, L.A., Delphia, C.M. and O'Neill, K.M. (2017) A dual role for farmlands: food security and pollinator conservation D. Gibson, ed. *J. Ecol.*, **105**, 890–899.

Cane, J.H. (2008) Breeding biologies, seed production and species-rich bee guilds of Cleome lutea and Cleome serrulata (Cleomaceae). *Plant Species Biology*, **23**, 152–158.

Carneiro, J.A.A., Rebouças, N.C., Ribeiro, R. de T.M., et al. (2018) Flora of Ceará, Brazil: Cleomaceae. *Rodriguésia*, **69**, 1659–1672.

Castro, T.C. De, Simões-gurgel, C., Ribeiro, I.G., Garcia, M., Coelho, P. and Albarello, N. (2014) Morphological aspects of fruits, seeds, seedlings and in vivo and in vitro germination of species of the genus *Cleome*. *J. Seed Sci.* **36**, 326–335.

Causier, B., Schwarz-Sommer, Z. and Davies, B. (2010) Floral organ identity: 20 years of ABCs. *Semin. Cell Dev. Biol.*, **21**, 73–79.

Cho, L.H., Yoon, J. and An, G. (2017) The control of flowering time by environmental factors. *Plant J.*, **90**, 708–719.

Cresswell, C.J., Cunningham, H.M., Wilcox, A. and Randall, N.P. (2018) What specific plant traits support ecosystem services such as pollination , bio - control and water quality protection in temperate climates? A systematic map. *Environ. Evid.*, 1–13.

Davies, B., Motte, P., Keck, E., Saedler, H., Sommer, H. and Schwarz-Sommer, Z. (1999) PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. *EMBO J.*, **18**, 4023–34.

Descamps, C., Quinet, M., Baijot, A. and Jacquemart, A.L. (2018) Temperature and water stress affect plant–pollinator interactions in Borago officinalis (Boraginaceae). *Ecol. Evol.*, **8**, 3443–3456.

Durand, B. and Durand, R. (1991) Sex determination and reproductive organ differentiation in Mercurialis. *Plant Sci.*, **80**, 49–65.

Estelle Jaligot, Adler S., Debladis E., Beulé T., Richaud F., Ilbert P., Jean Finnegan E., Rival A. (2011) Epigenetic imbalance and the floral developmental abnormality of the in vitro-regenerated oil palm Elaeis guineensis. *Ann. Bot.*, **108**, 1453–1462.

Ekpong, B. (2009) Effects of seed maturity, seed storage and pre-germination treatments on seed germination of *Cleome* (*Cleome gynandra* L.). *Scientia Horticulturae*, **119**, 236–240.

Feodorova, T. A., Voznesenskaya, E. V., Edwards, G.E. and Roalson, E.H. (2010) Biogeographic Patterns of Diversification and the Origins of *Cleome* (Cleomaceae). *Syst. Bot.*, **35**, 811–826.

Flora do Brasil (2021) Flora do Brasil 2020 - Algas, Fungos e Plantas. Available at: http://reflora.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/BemVindoConsultaPublicaConsultar.do?invalidatePageControlCounter=1&idsFilhosAlgas=%5B2%5D&idsFilhosFungos=%5B1%2C11%2C10%5D&lingua=&grupo=5&genero=Tarenaya&especie=hassleriana&autor=&nomeVer.

George Chuck, Meeley R., Irish E., Sakai H., Hake S. (2007) The maize tasselseed4 microRNA controls sex determination and meristem cell fate by targeting Tasselseed6/indeterminate spikelet1. *Nat. Genet.*, **39**, 1517–1521.

Garibaldi, L.A., Aizen, M.A., Cunningham, S. and Klein, A.M. (2009) Pollinator shortage and global crop yield. *Commun. Integr. Biol.*, **2**, 37–39.

Ghazoul, J. (2005) Buzziness as usual? Questioning the global pollination crisis. *Trends Ecol. Evol.*, **20**, 367–373.

Gomez Raboteaux, N.N. and Anderson, N.O. (2010) Germination of *Cleome hassleriana* and *Polanisia dodecandra* seed lots in response to light, temperature and stratification. *Res. J. Seed Sci.*, **3**, 1–17.

Guzella, T.S., Dey, S., Chelo, I.M., Pino-querido, A. and Veronica, F. (2017) Slower environmental change hinders adaptation from standing genetic variation. *Plos Genetics*, 1–29.

Hallmann, C.A., Sorg, M., Jongejans, E., et al. (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas E. G. Lamb, ed. *PLoS One*, **12**, e0185809.

Heller, S., Joshi, N.K., Leslie, T., Rajotte, E.G. and Biddinger, D.J. (2019) Diversified Floral Resource Plantings Support Bee Communities after Apple Bloom in Commercial Orchards. *Sci. Rep.*, **9**, 1–13.

Inda, L.A., Torrecilla, Æ.P., Catala, Æ.P. and Ruiz-zapata, T. (2008) Phylogeny of *Cleome* L. and its close relatives *Podandrogyne* Ducke and *Polanisia* Raf . (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. *Plant Systematics and Evolution*, **274**, 111–126.

Jack, T., Brockman, L.L. and Meyerowitz, E.M. (1992) The Homeotic Gene APETALA3 of Arabidopsis thaliana Encodes a MADS Box and Is Expressed in Petals and Stamens. *Cell*, **68**, 683–697.

Klein, A.M., Vaissière, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.A., Kremen, C. and Tscharntke, T. (2007) Importance of pollinators in changing landscapes for world crops. *Proc. R. Soc. B Biol. Sci.*, **274**, 303–313.

Koteyeva, N.K., Voznesenskaya, E. V, Roalson, E.H. and Edwards, G.E. (2011) Diversity in forms of C4 in the genus *Cleome* (Cleomaceae). *Ann Bot.*, 269–283.

Kozlov, K., Sokolkova, A., Lee, C.R., Ting, C.T., Schafleitner, R., Bishop-von Wettberg, E., Nuzhdin, S. and Samsonova, M. (2020) Dynamical climatic model for time to flowering in Vigna radiata. *BMC Plant Biol.*, **20**, 1–15.

Kremen, C. (2018) The value of pollinator species diversity. *Science (80-.).*, **359**, 741–742.

Lai, X., Vega-Léon, R., Hugouvieux, V., et al. (2021) The intervening domain is required for DNA-binding and functional identity of plant MADS transcription factors. *Nat. Commun. 2021 121*, **12**, 1–13.

Lai, Y.S., Shen, D., Zhang, W., et al. (2018) Temperature and photoperiod changes affect cucumber sex expression by different epigenetic regulations. *BMC Plant Biol.*, **18**, 1–13.

Larson, B.M.H. and Barrett, S.C.H. (2000) A comparative analysis of pollen limitation in flowering plants. *Biol. J. Linn. Soc.*, **69**, 503–520.

Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowmant, J.L. and Yanofsky, M.F. (2000) SHATTERPROOF MADS-box genes control dispersal in *Arabidopsis*. *Nature*, **404**, 766–770.

Lloyd, D.G. and Schoen, D.J. (1992) Self- and cross-fertilization in plants. I. Functional dimensions. *Int. J. Plant Sci.*, **153**, 358–369.

Luciano, R., Neto, S., Thomas, W.W., Regina, M., Barbosa, D.V. and Roalson, E.H. (2018) New combinations and taxonomic notes for Tarenaya (Cleomaceae). *Acta Bot. Bras.*, **32**, 540–545.

Machado, I., Cristina Lopes, A., Valentina Leite, A. and Virgíniade Brito Neves, C. (2006) Cleome spinosa (Capparaceae): polygamodioecy and pollination by bats in urban and

Caatinga areas, northeastern Brazil. *Bot. Jahrbücher für Syst. Pflanzengeschichte und Pflanzengeographie*, 69–82.

Marshall, D.M., Muhaidat, R., Brown, N.J., Liu, Z., Stanley, S., Griffiths, H., Sage, R.F. and Hibberd, J.M. (2007) *Cleome*, a genus closely related to *Arabidopsis*, contains species spanning a developmental progression from C3 to C4 photosynthesis. *Plant J.*, 886–896.

Marshman, J., Blay-Palmer, A. and Landman, K. (2019) Anthropocene Crisis: Climate Change, Pollinators, and Food Security. *Environments*, **6**, 22.

Matias, R., Oliveira, A.S. De, Furtado, M.T., Sá, T., Rodrigues, E.B., Oliveira, P.E. De and Consolaro, H. (2016) Atypical reproductive system of two Rubiaceae species: Dystilia with partial self-incompatibility in the brevistile morph? *Rodriguesia*, **67**, 357–368.

Mesgaran, M.B., Matzrafi, M. and Ohadi, S. (2021) Sex dimorphism in dioecious Palmer amaranth (*Amaranthus palmeri)* in response to water stress. *Planta*, **254**, 17.

Muyle A., Zemp N., Deschamps C., Mousset S., Widmer A., Marais G.A.B. (2012) Rapid de novo evolution of X chromosome dosage compensation in Silene latifolia, a plant with young sex chromosomes. *PLoS Biol.*, **10**.

Novais, S.M.A., Nunes, C.A., Santos, N.B., D'Amico, A.R., Fernandes, G.W., Quesada, M., Braga, R.F. and Neves, A.C.O. (2016) Effects of a possible pollinator crisis on food crop production in Brazil. *PLoS One*, **11**.

Nozzolillo, C., Treyvaud, V., Bily, A.C., Harris, C.S., Saleem, A., Andersen, ØM. and Jordheim, M. (2010) Novel aspects of the flowers and floral pigmentation of two *Cleome* species (Cleomaceae), *C . hassleriana* and *C . serrulata*. *Biochemical Systematics and Ecology*, **38**, 361–369.

Ollerton, J., Winfree, R. and Tarrant, S. (2011) How many flowering plants are pollinated by animals ? *Oikos,* **120**, 321–326.

Omondi, E.O., Debener, T., Linde, M., Abukutsa-Onyango, M., Dinssa, F.F. and Winkelmann, T. (2017) Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries M. Havey, ed. *Plant Breed.*, **136**, 578–589.

Patchell, M.J., Bolton, M.C., Mankowski, P. and Hall, J.C. (2011) Comparative Floral Development in Cleomaceae Reveals Two Distinct Pathways Leading to Monosymmetry. *Int. J. Plant Sci.*, **172**, 352–365.

Pawełkowicz, M.E., Skarzyńska, A., Pląder, W. and Przybecki, Z. (2019) Genetic and molecular bases of cucumber (*Cucumis sativus* L.) sex determination. *Mol. Breed.*, **39**, 1–27.

Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E. and Yanofsky, M.F. (2000) B and C floral organ identity functions require SEPALLATTA MADS-box genes. *Nature*, **405**, 200–203.

Popp, J., Pető, K. and Nagy, J. (2013) Pesticide productivity and food security. A review. *Agron. Sustain. Dev.*, **33**, 243–255.

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O. and Kunin, W.E. (2010) Global pollinator declines: Trends, impacts and drivers. *Trends Ecol. Evol.*, **25**, 345–353.

Prasifka, Jarrad R., Mallinger, R.E., Portlas, Z.M., Hulke, B.S., Fugate, K.K., Paradis, T., Hampton, M.E. and Carter, C.J. (2018) Using nectar-related traits to enhance crop-pollinator interactions. *Front. Plant Sci.*, **9**.

Prasifka, Jarrad R, Mallinger, R.E., Portlas, Z.M., Hulke, B.S., Fugate, K.K., Paradis, T., Hampton, M.E., Carter, C.J. and Campbell, A.J. (2018) Using Nectar-Related Traits to Enhance Crop-Pollinator Interactions. *Front Plant Sci,* **9**, 1–8.

Reeves, G., Singh, P., Rossberg, T.A., Sogbohossou, E.O.D., Schranz, M.E. and Hibberd, J.M. (2018) Natural variation within a species for traits underpinning C4 photosynthesis. *Plant Physiol.*, **177**, 504–512.

Riaz, S., Abid, R., Syed, A., Ali, S.A., Munir, I. and Qaiser, M. (2019) Morphology and seed protein profile for a new species of the genus *Cleome* L. (Cleomaceae) from Pakistan. *Protein Chem. to Proteomics*, 1–10.

Salzman, S., Crook, D., Calonje, M., Stevenson, D.W., Pierce, N.E. and Hopkins, R. (2021) Cycad-Weevil Pollination Symbiosis Is Characterized by Rapidly Evolving and Highly Specific Plant-Insect Chemical Communication. *Front. Plant Sci.*, **12**, 816.

Schilling, S., Pan, S., Kennedy, A. and Melzer, R. (2018) MADS-box genes and crop domestication: the jack of all traits. *J. Exp. Bot.*, **69**, 1447–1469.

Schoen, D.J., Johnston, M.O., L'Heureux, A.-M. and Marsolais, J. V. (1997) Evolutionary History of the Mating System in *Amsinckia* (Boraginaceae). *Evolution (N. Y).*, **51**, 1090–1099.

Schranz, M.E. and Mitchell-olds, T. (2006) Independent Ancient Polyploidy Events in the Sister Families Brassicaceae and Cleomaceae. *The Plant Cell*, **18**, 1152–1165.

Shao, F., Wang, S., Liu, Z., Hong, R. and Zhao, T. (2020) Pollination, fertilization, and embryo development in southern China fresh-eating Jujube. *HortScience*, **55**, 1315–1322.

Sidhu, C.S. and Joshi, N.K. (2016) Establishing wildflower pollinator habitats in agricultural farmland to provide multiple ecosystem services. *Front. Plant Sci.*, **7**.

Simões, C. (2006) In vitro propagation of *Cleome spinosa* (Capparaceae) using explants from nursery- grown seedlings and ...

Sobrevila, C. and Arroyo, M.T.K. (1982) Breeding systems in a montane tropical cloud forest in Venezuela. *Plant Syst. Evol. 1982 1401*, **140**, 19–37.

Soltis, P.S., Folk, R.A. and Soltis, D.E. (2019) Darwin review: angiosperm phylogeny and evolutionary radiations. *Proc. R. Soc. B Biol. Sci.*, **286**.

Stevens, P.F. Angiosperm Phylogeny Website. Version 14, July 2017 [and more or less continuously updated since].

Takebayashi, N. and Morrell, P.L. (2001) Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am. J. Bot.*, **88**, 1143–1150.

Telgmann-Rauber, A., Jamsari, A., Kinney, M.S., Pires, J.C. and Jung, C. (2007) Genetic and physical maps around the sex-determining M -locus of the dioecious plant asparagus. *Mol. Genet. Genomics 2007 2783*, **278**, 221–234.

Vinícius-Silva, R., Parma, D. de F., Tostes, R.B., et al. (2017) Importance of bees in pollination of *Solanum lycopersicum* L. (Solanaceae) in open-field of the Southeast of Minas Gerais State, Brazil. *Hoehnea*, **44**, 349–360.

Voznesenskaya, E. V., Koteyeva, N.K., Chuong, S.D.X., Ivanova, A.N., Barroca, J., Craven, L.A. and Edwards, G.E. (2007) Physiological, anatomical and biochemical characterisation of photosynthetic types in genus *Cleome* (Cleomaceae). *Funct. Plant Biol.*, **34**, 247.

Wang, L., Czedik-eysenberg, A., Mertz, R.A., et al. (2014) Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nat. Biotechnol.*, **32**, 1158–1165.

Williams, N.M., Ward, K.L., Pope, N., et al. (2015) Native wildflower plantings support wild bee abundance and diversity in agricultural landscapes across the United States. *Ecol. Appl.*, **25**, 2119–2131.

Wratten, S.D., Gillespie, M., Decourtye, A., Mader, E. and Desneux, N. (2012) Pollinator habitat enhancement: Benefits to other ecosystem services. *Agric. Ecosyst. Environ.*, **159**, 112–122.

She X., Xu X., Fedotov A., Kelly W.G., Maine, L.M. (2009) Regulation of heterochromatin assembly on unpaired chromosomes during *Caenorhabditis elegan*s meiosis by components of a small RNA-mediated pathway. *PLoS Genet.*, **5**.

Zapata, T.R. and Arroyo, M.T.K. (1978) Plant Reproductive Ecology of a Secondary Deciduous Tropical Forest in Venezuela. *Biotropica*, **10**, 221.

Zohoungbogbo, H.P.F., Houdegbe, C.A., Sogbohossou, D.E.O., Tossou, M.G., Maundu, P., Schranz, E.M., Deynze, A. Van, Zoundjihekpon, J. and Achigan-Dako, E.G. (2018) Andromonoecy in *Gynandropsis gynandra* (L.) Briq. (Cleomaceae) and effects on fruit and seed production. *Genet. Resour. Crop Evol.*, **65**, 2231–2239.

SUPPLEMENTARY MATERIAL

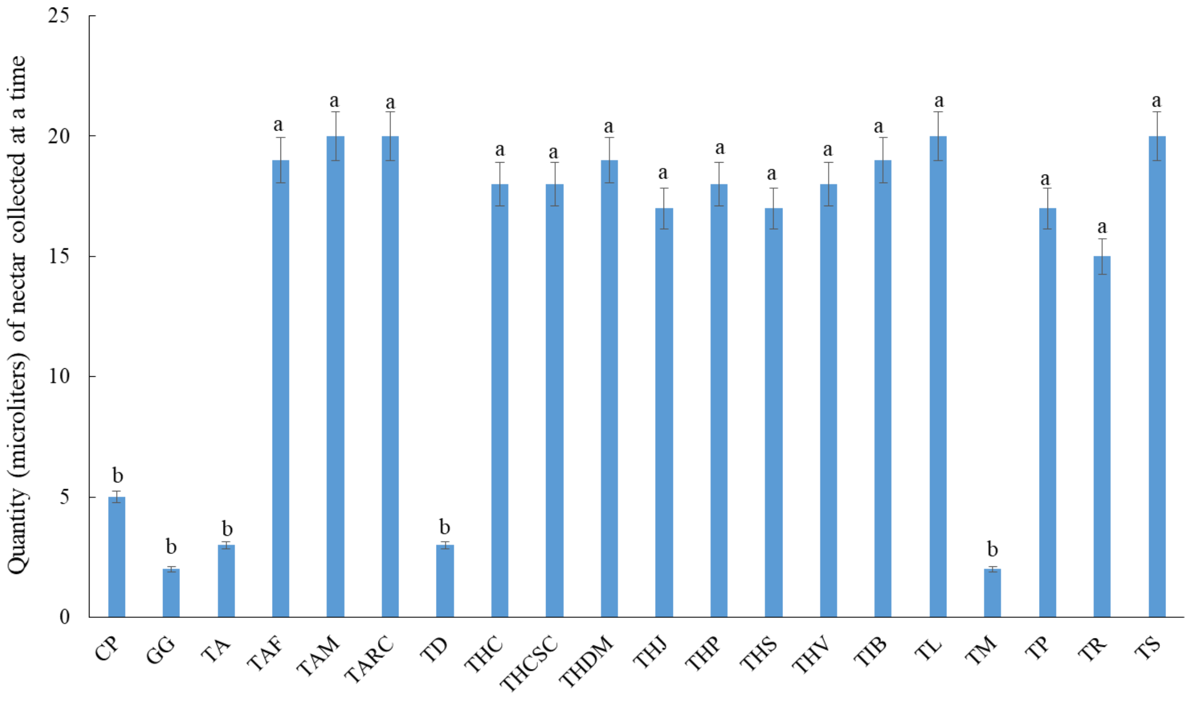


Figure 1: Amount of nectar collected at one time per species.

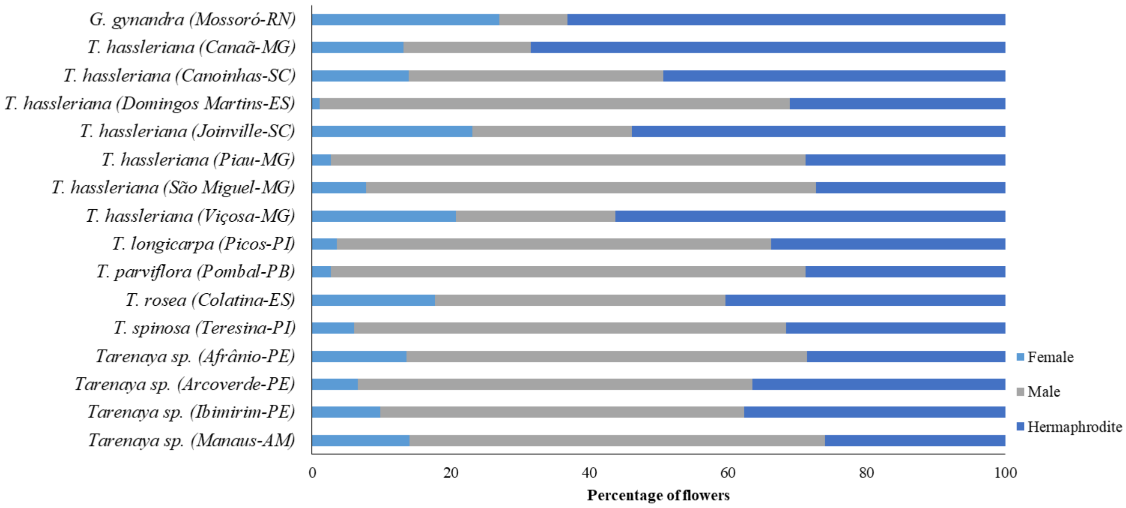


Figure 2. Percentage of flowers produced by polygamous individuals of 16 Cleomaceae species (n = 150).

Figure 3: Flowers of *Tarenaya diffusa*. A. Chasmogamous flower. B-D. Cleistogamous flowers.

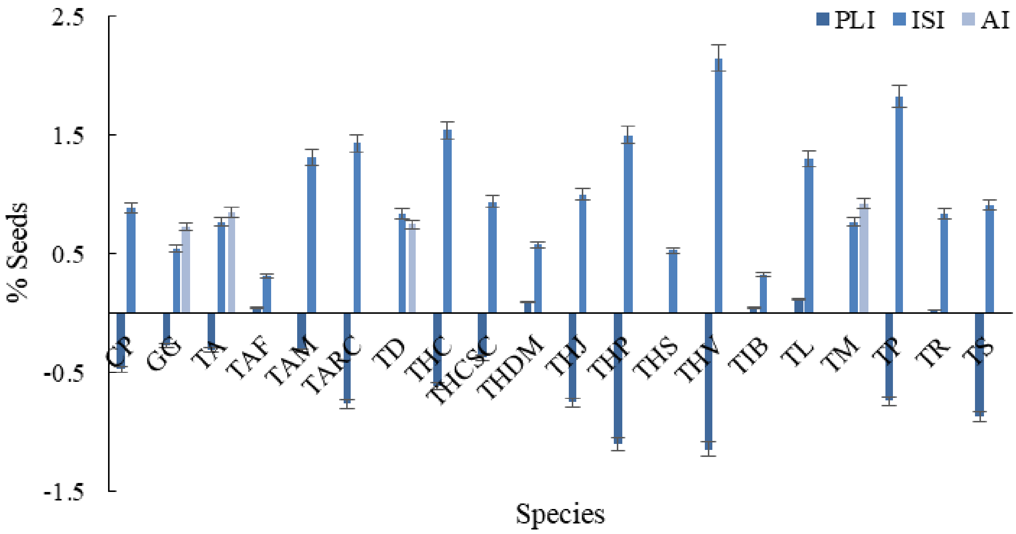


Figure 4: Percentage of seeds produced according to each variable. Pollen limitation (PLI), according Larson and Barrett, (2000), index of self-incompatibility (ISI) and self-pollination index (AI) according Lloyd and Schoen, (1992).

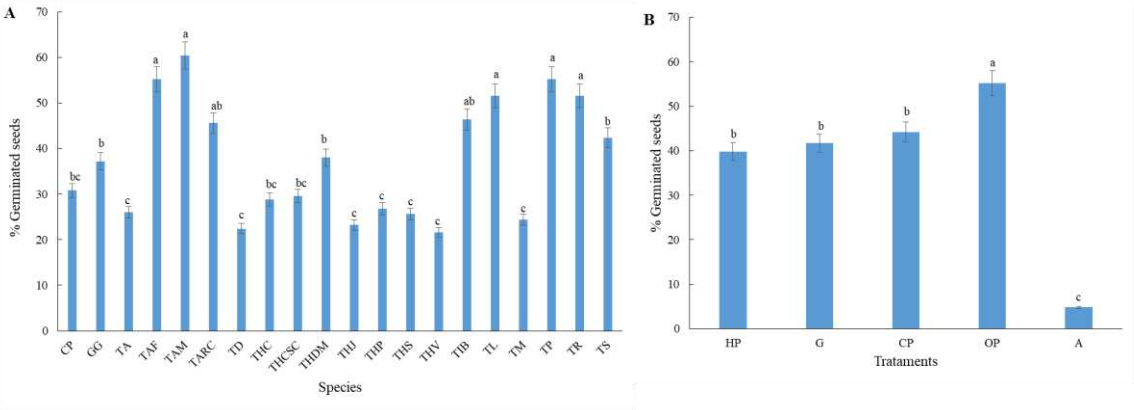


Figure 5: Germinated seeds by pollination treatments for Cleomaceae species. (**A**)- Average seed germinated by species. (**B**)-Average germinated seeds per treatment. Letters above individual bars indicate significant groupings according to Tukey's Test, n=5.

Table 1: Cleomaceae species used in this study followed by their identification code (ID), city and state of collection, herbarium voucher and NCBI accession number for ITS partial gene sequences.

Species	**ID**	**City**	**State**	**Voucher**
Cleoserrata paludosa	CP	Belém	PA	VIC 51788
Gynandropsis gynandra	GG	Mossoró	RN	VIC 51789
Tarenaya aculeata	TA	Feira de Santana	BA	VIC 51791
Tarenaya diffusa	TD	Feira de Santana	BA	VIC 51793
Tarenaya hassleriana	THC	Canaã	MG	VIC 51795
Tarenaya hassleriana	THCSC	Canoinhas	SC	VIC 51900
Tarenaya hassleriana	THDM	Domingos Martins	ES	VIC 51901
Tarenaya hassleriana	THSM	São Miguel	MG	VIC 51796
Tarenaya hassleriana	THJ	Joinville	SC	VIC 51802
Tarenaya hassleriana	THP	Piau	MG	VIC 51902
Tarenaya hassleriana	THV	Viçosa	MG	VIC 51794
Tarenaya longicarpa	TL	Picos	PI	VIC 51799
Tarenaya microcarpa	TM	Belém	PA	VIC 51797
Tarenaya parviflora	TP	Pombal	PB	VIC 51798
Tarenaya rosea	TR	Colatina	ES	VIC 51800
Tarenaya longicarpa	TAM	Manaus	AM	VIC 51803
Tarenaya longicarpa	TAF	Afrânio	PE	VIC 51903
Tarenaya longicarpa	TIB	Ibimirim	PE	VIC 51904
Tarenaya longicarpa	TARC	Arcoverde	PE	VIC 51905
Tarenaya spinosa	TS	Teresina	PI	VIC 51804

Table 2: Results of the principal component analysis (PCA). The table shows the values of the scores obtained for each independent variable for each factor.

Variable	**Factor 1**	**Factor 2**	**Factor 3**
Andromonocy	0.252358	-0.967435	-0.019638
Hermaphrodite	0.898350	0.328024	0.292176
Poligamous	-0.939438	0.234299	-0.250119
Nectar	-0.893430	-0.357153	0.272440
Flower size	-0.952333	0.157007	0.261554

CHAPTER 6

A family with great possibilities: Is Cleomaceae a new model family?

Daniele F. Parma[1], Marcelo G. M. V. Vaz[1], Wagner L. Araújo[1], M. Eric Schranz[2], Andreas P. M. Weber[3] & Adriano Nunes-Nesi[1]*

[1]Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.
[2]Biosystematics Group, Wageningen University, 6708 PB Wageningen, The Netherlands.
[3]Institute for Plant Biochemistry Biology, Cluster of Excellence on Plant Science CEPLAS, Heinrich Heine University, Düsseldorf, Germany.

*Corresponding author:
Adriano Nunes-Nesi
Departamento de Biologia Vegetal,
Universidade Federal de Viçosa,
36570-900 Viçosa, Minas Gerais, Brazil
Phone: +55-31-3612-5357
Email: nunesnesi@ufv.br

ABSTRACT

The global population grows in number and the demand for food is likely to increase considerably in the near future. Thus, many researches aim to understand how plants adapt quickly to environmental factors and how this knowledge can be used in agriculture. In this sense, it is extremely important to understand how we can increase productivity in increasingly smaller areas. Most important, attention must be paid on how to cope with the climate changes that are already happening and with the extinction and reduction of pollinating agents that are essential in some crops. Therefore, Cleomaceae is emerging as a promising family to investigate a wide range of phenomena, such as C_4 photosynthesis and floral diversity. This family is sister to Brassicaceae, which can facilitate the transfer of knowledge from the plant model *Arabidopsis thaliana*. Investigations in C_4 photosynthesis are instrumental in untangling the origin and maintenance of this ecologically relevant trait. In turn, examination of floral symmetry in Cleomaceae provides an opportunity to compare monosymmetric flowers with the polysymmetric flowers generated by Brassicaceae. Remarkable floral diversity of Cleomaceae is an underutilized system to address the evolution of features associated with the important phenomenon of pollination. In addition, some species (*e.g. Gynandropsis gynandra*) can be consumed as vegetables. Furthemore, the family have the potential in the development of drugs for various diseases and can be used even as insecticides. Here, we compile the various opportunities for advancing studies in several areas with Cleomaceae, that although promising, functional techniques are lacking for elucidating this diversity.

Keywords: Brassicales, Cleomaceae, photosynthesis, hercogamy.

INTRODUCTION

Cleomaceae is a small family of the order Brassicales that has 26 formally described genera, comprising about 270 species (Bayat et al., 2018; Stevens, 2001; Soares-Neto et al., 2020a). Thisfamily has a worldwide cosmopolitan distribution (Stevens, 2001), originating in central Asia (Feodorova et al., 2010), andwith a center of diversity in the Americas. Cleomaceaeis characterized by having a diverse habit (herbaceous to shrub); monosymmetric flowers, with a ground plan of four sepals; four petals; generally six stamens and a bicarpellate gynoecium (Stevens, 2001). The flowers are diverse in shape, size and color, as well as in the presence and size of the nectariferous disc, characteristics that influence the type of pollination (Stevens, 2001).

The family is phylogenetically close to Brassicaceae (sister families), in which the *Arabidopsis thaliana* model plant is inserted (Hall et al., 2002; Stevens, 2001). Thus, Cleomaceae, which has a great physiological, anatomical and morphological diversity, has been the focus of important ecological and evolutionary inquiries. These studies examined floral morphology and development (Nozzolillo et al., 2010; Patchell et al., 2014), the evolution of C_4 photosynthesis (Koteyeva et al., 2011; Marshall et al., 2007; Voznesenskaya et al., 2007), pollination biology (Cane, 2008; I. Machado et al., 2006) and comparative genomics (Bergh et al., 2014; Cheng et al., 2013; Schranz and Mitchell-olds, 2006). Although the family is important in studies of several areas, few species have been studied in detail. In this regard much of what we know about the family is based on the species *Gynandropsis gynandra*, *Tarenaya hassleriana* and *T. spinosa.* But by far the most studied species is the African and Asian crop *G. gynandra*. The lack of studies with other members of Cleomaceae may be due to the few basic works developed, such as: (i) floristic surveys and taxonomy - collection, ecological distribution and identification, (ii) seed germination, considering that some species have dormancy and low viability, (iii) karyotyping, and (iv) pollination studies. In this regard, studies in these areas have been increasing in recent years, as well as their scope for New World species, which until then were little explored.

Thus, this family has great potential to allow advances in several fields of science,. Accordingly, in this review we have brought a compilation of studies with the family in different areas and we try to relate this different information to identify evolutionary characters in Clemoceae. In this sense, we hope that this work will contribute to the reflective aspect, demonstrating the importance of communicating the different fields and data available, as well as being a reference as a starting point to direct work with this family.

HISTORY AND CLASSIFICATION

Taxonomic studies with Cleomaceae are relatively frequente. In this sense, we can mention the description of some genera that were included in the Capparaceae subfamily, Cleomoideae. *Cleomella*, *Peritoma*, *Polanisia* (De Candolle 1824), *Dactylaena*, *Physostemon* (Eichler 1865; Pax and Hoffman 1936), *Podandrogyne* (Ducke 1930), *Haptocarpum*, *Cristatella*, *Gynandropsis*, *Isomeris*, *Justago*, *Tetratelia* (Pax and Hoffman 1936; Kers 2003), *Oxystilis* and *Wislizenia* (Iltis 1957, Keller 1979) are some examples. Eichler (1865) recognized only four genera within the Cleomae tribe (*Cleome* s.l., *Gynandropsis*, *Dactylaena* and *Physostemon*) and the genus *Cleome* s.l. was separated into three series based on habit. Subsequently, based on characters related to the shape of the fruit and floral buds, the presence or absence of large foliar bracts, aculeate stipules and the type of pubescence of the plant, Heiborn (1931) recognized seven groups within the series *Cleome* s.l. created by Eichler (1865). These studies provided the basis for the study by Pax and Hoffmann (1936), which became the most comprehensive worldwide systematic treatment of *Cleome* and its close relatives. Accordingly, *Cleome* s.l. (Cleomoideae) was divided into two supergroups (Old Word and New Word), two sections, four subsections and six series, based on morphological characters, such as leaf division, habit and stamen number (Pax and Hoffman 1936). However, this classification was largely based on characters that were not reliable taxonomically because of their great variability within the genus (Sanchez-Acebo, 2005). Iltis (1952), in turn, considering the characteristics of the seed, nectariferous disk, raceme morphology, floral aestivation and leaves, divided New World *Cleome* into two subgenera, 10 sections, and 17 series (Inda et al., 2008; Sanchez-Acebo, 2005).

Phylogenetic studies with the Cleomaceae family started by trying to understand the circumscription of Capparaceae and Brassicaceae, since Cleomoideae (subfamily of Capparaceae) positioned itself as closer to Brassicaceae than Capparaceae (Hall et al., 2002; 2004; Hall 2008; Iltis et al. 2011). Thus, these studies identified that Cleomoideae, Capparoideae, and Brassicoideae form well-supported monophyletic and distinct clades. In that sense, some authors (*e.g.* Airy Shaw 1965; Hutchinson 1967; Hall et al., 2002) recommended to recognize these subfamilies as three separate families, Capparaceae, Cleomaceae, and

Brassicaceae. However, it was only in 2011 that Iltis et al., (2011) raised Cleomoideae to the family level, based on previous works of molecular phylogeny and morphological characterization. On this subject, Cleomaceae differs from Capparaceae and Brassicaceae by their palmately compound leaves, capsular fruits lacking a septum, and seeds ornamented with a testa that has a pronounced invagination (Hall et al., 2002; Iltis et al., 2011).

After the recognition of Cleomaceae as a family, the focus became understanding the real circumscription of the genus *Cleome*, since it was very specious (about 200 species) and proved to be polyphyletic in molecular phylogeny studies (*e.g.* Hall, 2008; Patchell et al., 2014; Sanchez-Acebo, 2005). The most complete work for this family, based on molecular phylogeny, aimed to understand the genus *Cleome* s.s., sampling the type species that until then had not been analyzed in a phylogenetic context, and using five molecular markers (ITS, matk, ndhF, ycF and rps3) and three genomes (nuclear, chloroplast, mitochondrial) (Patchell et al., 2014). After this study, several others emerged to delimit the genus genera of this family and make them monophyletic (*e.g* Barrett et al., 2017; Luciano et al., 2018; Roalson et al., 2015, 2017; Thulin et al., 2017). Thus, the family currently has 26 genera and about 270 species (Bayat et al., 2018; Stevens, 2001; Soares-Neto et al., 2020a) (Fig. 1). However, it is worth mentioning that the number of genera and species tends to increase with new research in the area (*e.g.* Soares-Neto et al. 2020b).

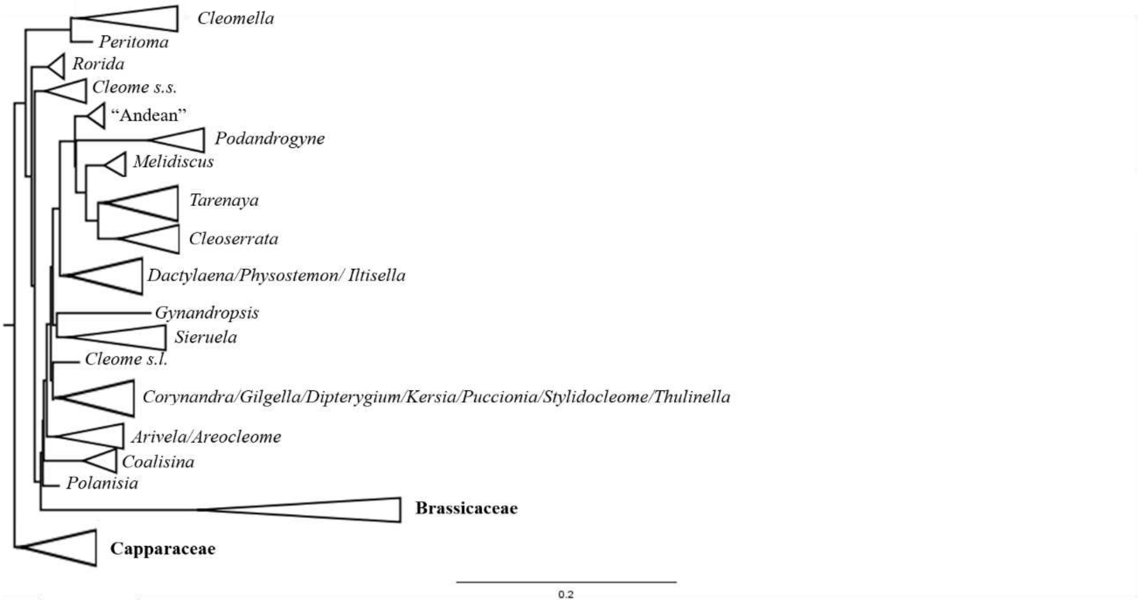


Figure 1: Twenty-six genera of Cleomaceae have been identified. Traditionally circumscribed Cleome are distributed across multiple clades.

ASPECTS OF MORPHOLOGICAL EVOLUTION WITHIN CLEOMACEAE

In Cleomaceae species few studies attempted to associate morphology with molecular phylogeny (*e.g.* Sanchez-Acebo, 2005). This association of traits is essential for recognizing groups of cleomes in order to trace evolutionary aspects. It is worth mentioning that all morphological characteristics (vegetative and reproductive) are important to be evaluated. However, the reproductive traits better represent the evolutionary history of the group. In this sense, seed morphology, pollen, floral symmetry, number of stamens and petal shape are some of the reproductive characteristics usually evaluated (*e.g.* Sanchez-Acebo, 2005).

In this respect, it has been demonstrated that the pollen of Cleomaceae is tricolporate and highly variable in ornamentation, ranging from echinate to reticulate, rugulate or verrucate-psilate (Sanchez-Acebo, 2005). Thust, there is a tendency to separate groups in Cleomaceae based on the characteristic of the pollen. Accordingly, the most basal groups of the family have reticulate or rugulate pollen and the most echinate-derived pollen groups (Sanchez-Acebo, 2005). In this aspect, based on studies with the theme, members of the old world have pollen reticulate or rugulate (*e.g.* North American group - Al-Shehbaz, (1973), African cluster, *Polanisia* sp., *G. gynandra* - Sanchez-Acebo, (2005), *C. viscosa*, *C. karachiriensis* - Riaz et al., (2019)), an exception would be *Physostemon* (Ruiz-Zapata and Enrech 1997). Pollen morphology of the echinate ornamentation type is present in *Tarenaya*, Andean group, *Podandrogyne* sp., *Cleome stylosa*, *C. lechleri*, *C. anomala* and *C. moritziana* (Sanchez-Acebo, 2005). This suggests that echinate pollen may be derived from reticulate or rugulate pollen (Sanchez-Acebo, 2005). However, unlike the other mentioned species mentioned, *Melidiscus giganteus* have pollen verrucate (Sanchez-Acebo, 2005) and *C. brachycarpa* has spinulose ornamentation (Riaz et al., 2019).

Regarding the seed morphology, Cleomaceae has several types of ornamentation, ranging from smooth (e.g. *Cleoserrata paludosa* – Sanchez-Acebo, 2005) to verrucose seeds. The *Tarenaya* species, such as *T. aculeata*, *T. crenopetala*, *T. dendroides*, *T. diffusa*, *T. eosina*, *T. hassleriana*, *T. horrida*, *T. microcarpa*, *T. parviflora*, *T. regnelli*, *T. rosea*, *T. spinosa*, *T. titubans*, *T. trachycarpa* and *T. virens* have ornamentation of the seeds as reticulate, reticulate-areolated with transversal, ribbed and alveolar ridges, in addition to the presence of aryl in the seeds of some species (Costa-Silva, 2000). The seeds of *Melidiscus giganteus* is smooth (Sanchez-Acebo, 2005). In general, the Old World species tend to have the seed cleft open, but it may be so narrow as to appear closed. In turn, in *Tarenaya* spp. and *M. giganteus*, New World group, have seed cleft closed for most of its length by a thin to thick membrane. In this sense, although it is not possible to trace a pattern of ornamentation of pollen and seeds as a synapormophy among Cleomaceae, due to the several exceptions that occur, it is possible to identify, only with these characteristics that Old World species mostly basal to the New World species (Sanchez-Acebo, 2005), which is congruent with molecular phylogeny (Fig. 1).

The floral morphology of Cleomaceae is quite diverse, being mostly zygomorphic in symmetry. This floral monosymmetry arises through upwards curvature of corolla and androecium whorls, which may be complemented by shape, size and color differences between adaxial and abaxial petals as well as variation in nectar gland shape (Kers, 2003; Patchell et al., 2011). Regarding the number of stamens, most genera of Cleomaceae have about six stamens (*e.g. Cleomella*, *Cleoserrata*, *Gynandropsis*, *Mitostylis*, *Tarenaya*). In turn, some more basal groups such as *Kersia* (up to 35 stamens) and *Corynandra* (up to 250 stamens) can have numerous stamens (Bayat et al., 2018). Apparently, there is a tendency to reduce the number of stamens within the family, as seen for angiosperms in general. However, similarly to what was observed previously for pollen and seed morphology, there are exceptions, so it is possible to find species that belong to the Old World group, such as *Cleome viscosa* (10-20 stamens), *C. karachiriensis* (6-8 stamens) and *C. brachycarpa* (6 stamens) that have few stamens (Riaz et al., 2019). In this sense, based on these characteristics it is also difficult to trace a synapormophy for the family.

FLORAL DIVERSITY IN CLEOMACEAE

The floral diversity observed in Cleomaceae allows for various pollination strategies, from spontaneous self-pollination to biotic cross-pollination. In this sense, for this last strategy, Cleomaceae in general produce large amounts of nectar and pollen. Some species of the family have nocturnal anthesis and are pollinated by bats and moths (*e.g. Tarenaya spinosa* - Machado et al., 2006), others in turn are pollinated by bees and butterflies (Cane, 2008; Zohoungbogbo et al., 2018). In addition, other mechanisms like hercogamy, physical barrier, (*e.g. G. gynandra* - Zohoungbogbo et al., 2018) and dicogamy, temporal barrier (*e.g. T. hassleriana*) are present in hermaphrodite species as an escape from self pollination, leading to increases genetic variability. The hercogamy – dysthilia, probably appeared more than once in Cleomaceae, since some groups do not have dysthilia (*e.g. Cleome brachycarpa*, *C. violacea* and *C. karachiensis* has homostylia – Riaz et al., 2019). Studies exploring floral morphology and reproductive biology in Cleomaceae are scarce and, limited to accesses from Africa (Omondi et al., 2017; Zohoungbogbo et al., 2018), North America (Cane, 2008) and Northeast of Brazil (Machado et al., 2006) in specific spots.

The diversity observed in Cleomaceae might lead to a better understanding in a pollination context and floral morphology. It is worth mentioning that the decline of pollinators has been concerning worldwide, with increasing evidence that pollinators are in crisis (Hallmann et al., 2017; Marshman et al., 2019; Novais et al., 2016) and that Cleomaceae has the potential to be a model for studies on the development of varieties and hybrids with improved floral traits that reduce dependence on pollinators or the production of nectar or pollen.

In this sense, studies with genes of family of MADS-domain transcription factors might allow us to discover the transcription factors (TFs) responsible for the diversity in floral morphology observed in Cleomaceae, as well as to understand the different types (sex) of flowers. It was demonstrated that *T. hassleriana* (Cleomaceae) has two copies of the floral B-class gene PI, one similar to that found in other angiosperms, while the second more related to that observed for Brassicaceae members (de Bruijn et al., 2018). In addition to differences between families, it is also possible to observe differences in the number of copies of some genes, as well as their expression within the same family. *T. hassleriana* has two AP3 paralogs originate from a recent tandem duplication and they are highly similar while *G. gynandra* do not present this tandem duplication (de Bruijn et al., 2018). In turn, to understand the dysthilia observed in the family, much progress is still missing. There is evidence of supergene control of the morphological and physiological components of heterostyly, but virtually nothing is known (Barrett, 2002). Thus, understanding the factors behind the floral diversity will allow us to promote species for cross-pollination or even to avoid it (self-pollination).

Accordingly, the development of crop varieties with specific nectar or nectar-related traits to attract and retain pollinating insects is an appealing strategy to address needs of agriculture and pollinators preservation (Prasifka et al., 2018). Thus, Cleomaceae emerges as a promising group for improving crop yields (promoting cross-pollination or not) and offering resources for bees. The family are attractive to both managed and wild pollinators, but because of selection for self-fertility, they are sometimes considered to have a low-to-moderate dependence on biotic pollinators.

GENETIC CHARACTERIZATION

While patterns of diversification in Cleomaceae have been well characterized with the use of molecular data (Hall, 2008; Patchell et al., 2014), other characters (*e.g.* micromorphological, physiological and genetic) have been explored to a limited extent. Among the available options, cytological data represent a potential valuable source of evolutionary and taxonomic information, once that it can provide important evidence of reproductive dynamics across different groups of angiosperms (Levin, 2002; Sanchez-Acebo, 2005). In this sense, cytological reports in Cleomaceae exist only in some species (Inda et al., 2008; Omondi et al., 2017; Subramania and Susheela, 1988). Besides this record, no additional information exists about ploidy number or karyotype variation for its remaining taxa. The size of the genome so far has been determined only for the species *G. gynandra*, with 2C = 2.31-2.45 pg (Omondi et al., 2017) and, more recently for Cleomaceae species occurring in Brazil, such as *Tarenaya* species (*T. hassleriana* with 2C = 0.55 pg, *T. diffusa* with 2C = 0.59 pg, *T. aculeata* and *T. microcarpa* with 2C = 0.66 pg, *T. siliculifera* with 2C = 0.77 and *T. spinosa, T. longicarpa, T. rosea, T. parviflora* with 2C = 1.30 pg), *Cleoserrata paludosa* 2C = 1.08, *G. gynandra* 2C = 2.20 (Parma et al. *submitted*).

The chromosomal number of Cleomaceae species has long been determined (eg. Ammal, 1933; Subramania and Susheela, 1988). However, there are few specific studies with the family for this purpose, some of the available data comes from more comprehensive studies such as characterization of angiosperms or Capparidaceae (family to which Cleomaceae was inserted). Thus, it appears that there is a tendency for most species in the family to be diploid, with some exceptions tetraploid groups (*e.g. C. chellidonii* – Subramania and Susheela, 1988; some *Tarenaya* species – Inda et al., 2008; *G. gynandra* – Omondi et al., 2017). In addition, the chromosome number tends to be higher in the most basal species groups (Old World), reaching x=29 em *Podandrogyne macrophylla*, *C. pilosa*, *C. stylosa*, *C. lechleri*, *C. anomala* and *C. moritziana* (Inda et al., 2008). Other groups, in turn, have an intermediate chromosome number to most species in the New World group, such as x=20 *C. multicalis* (Inda et al., 2008), x=16, 17 *C. chelidonii* (Subramania and Susheela, 1988), *G. gynandra* (Inda et al., 2008; Omondi et al., 2017); *Cleomella serrulata*, *Cleomella lutea* and *Melidiscus giganteus* (Inda et al., 2008). In turn, for the most derived species and with some exceptions from Old World groups, the chromosome number varies from 8-12, among these species, we highlight *C. viscosa*, *C. monophylla*, *C. tenella*, *C. felina*, *C.aspera* (Subramania and Susheela, 1988), *C. viscosa* (Ammal, 1933; Inda et al., 2008), *C. cordobensis*, *C. stenophylla*, *C. monohylla*, *C.*

rutidosperma, *C. violacea*, *Polanisia dodecandra*, *T. crenopetala*, *T. diffusa*, *T. microcarpa*, *T. boliviensis*, *T. hassleriana*, *T. longicarpa*, *T. werdemanii*, *T. titubans*, *T. spinosa*, *T. pernambucensis*, *T. parviflora*, *T. torticarpa*, *T. chapalensis*, *T. rosea*, *T. tucumanensis* and *T. aculeata* (Inda et al., 2008).

EVOLUTION OF C_4 PHOTOSYNTHESIS IN CLEOMACEAE

The C_4 photosynthetic mechanism emerged more than 60 times independently in 19 different families (Sage, 2004). The evolution of the C_4 photosynthetic mechanism arose from a C_3 plant, and the C_2 or intermediate C_3-C_4 mechanism would represent a transition between the two mechanisms (C_3 and C_4). Thus, it is known that there are some necessary steps for the development of C_4 photosynthesis, among them we can mention: (i) genetic precondintioning, (ii) anatomical preconditioning, (iii) enhancement of bundle sheath organelles, (iv) photorespiratory CO_2 pump glycine decarboxylase to bundle sheath cells, (v) enhancement of phophoenolpyruvate carboxylase activity, (vi) establishment of C_4 cycle and (vii) optimization of the C_4 syndrome (Sage, 2004; Gowik and Westhoff, 2011). In this sense, many efforts have been made to understand how this transition between mechanisms occurs, since the C_4 mechanism enables greater biomass production with less water use, compared to C_3 plants (Lundgren et al., 2014). The presence of C_4 enzymes in C_3 species suggest that it does not need to envolve *de novo* and so likely facilitated the recurrent evolution of the C_4 pathway (Reyna-llorens et al., 2017). However, it is still unclear how genes that have been co-opted into C_4 photosynthesis are regulated in C_3 species and to elucidate the exact role of each protein in C_3 plants (Brown et al., 2005).

The Cleomaceae family has representatives with the three different photosynthetic mechanisms, that together with its phylogenetic proximity to Brassicaceae, makes it a interesting model family to understand the important steps to acquire these characteristics. Moreover, it should be noted that in addition to being family-sisters, Brassicaceae and Cleomaceae, its genomes demonstrate significant synteny (Schranz and Mitchell-olds, 2006). Within the last two decades, studies using Cleomaceae family members with these objectives have been carried out In the last two decades, studies using Cleomaceae family members with these objectives have been carried out, however there are still few species physiologically characterized (only 15) (Marshall et al., 2007; Reeves et al., 2018; Voznesenskaya et al., 2007). Therefore, we can mention the work that started photosynthetic mechanism in the family, Imbamba and Tieszen (1977), Krenzer et al., (1975), Raghavendra and Das (1978) and Sankhla et al., (1975) who identified *Gynandropsis gynandra* with the C_4 photosynthetic mechanism and the studies of Rajendrudu and Rama Das (1982) that, in addition to confirming that *G. gynandra* has C_4 metabolism, evidenced that *Sieruela monophylla*, *Cleome aspera* and *C. viscosa* are C_3 species.

The next studies to identify the type of photosynthetic mechanism of Cleomaceae species took place in 2007, with Marshall et al., (2007) and Voznesenskaya et al., (2007) and, later in 2010 with Feodorova et al., (2010) and 2018 with Reeves et al., (2018). In this scenario, some species and different accessions of the family were described physiologically (Table 1). Thus, three species were characterized by presenting the photosynthetic metabolism C_4 (*Areocleome oxalidea*, *Coalisina angustifolia* and *G. gynandra*), one with the intermediate metabolism C_3-C_4 (*Coalisina paradoxa*) and the rest C_3 (Fig. 2). Although, it is noteworthy that in the C_3 species a flexibility was identified as to the development of C_3-C_4 photosynthesis (see Marshall et al., 2007), which may reflect ecological, genomic and anatomical factors that facilitate the acquisition of novel traits (Christin et al., 2013, 2012, 2011; Marshall et al., 2007; Mckown and Dengler, 2007; Sage, 2001). In addition, Reeves et al., (2018) demonstrated, when working with nine accesses by *G. gynandra,* that this species exhibit natural variation in key characteristics of C_4 photosynthesis. As consequence, the natural variation observed in *G. gynandra* can help to delineate genomic regions associated with C_4 trait variation (Reeves et al., 2018).

The C_4 mechanism identified in Cleomaceae species has different origins, which can be evidenced by both molecular phylogeny (Patchell et al., 2014) and different types of Kranz anatomy (*e.g. A. oxalidea* and *G. gynandra* have atriplicoid and *C. agustifolia*, angustifolioid Kranz anatomy) (Koteyeva et al., 2011). However, all three species with C_4 metabolism are of the NAD dependent malic enzyme (NAD-ME) subtype (Koteyeva et al., 2011). Aiming to characterize the species of the family regarding the photosynthetic mechanism, Voznesenskaya et al., (2007) analyzed 238 leaf materials (about 150 species) for isotopic carbon discrimination, found that values ranging from −12.3 ‰ to −33.9 ‰ and raised the hypothesis of three more origins of C_4 photosynthesis in the family (*Tarenaya siliculifera*, *Sieruela allamanii* and *S. gallaensis*). It is worth mentioning that isotopic discrimination is a tool that can assist in the identification of species with C_3 and C_4 metabolism. However, to identify species with the intermediate C_3-C_4 metabolism, gas exchange analyzes are necessary (Schluter and Weber, 2016). In this sense, there has been little research on leaf anatomy and the types of photosynthesis (by gas exchange) among members of the family Cleomaceae (*e.g.* Marshall et al. 2007; Voznesenskaya et al., 2007; Reeves et al. 2018).

So far, there has been talk about basic studies with the family (*e.g.* physiological characterization), which somehow support the more refined ones, such as comparing genomes (*e.g.* Bergh et al., 2014; Cheng et al., 2013; Külahoglu et al., 2014; Williams et al., 2016). In this sense, when comparing *G. gynandra* (C_4) and *T. hassleriana* (C_3), it was verified that both species share the Th-a WGD event, and there is no significant difference in retained gene copy number between them (Bergh et al., 2014), suggesting that differences in the timing and quantity of gene expression are responsible for the transition to C_4, and not the presence of additional gene copies (Bayat et al., 2018; Bergh et al., 2014). Moreover, gene expression atlas was generated across multiple developmental stages and tissue types in both species (Külahoglu et al., 2014), which can subsidize subsequent analyses.

Despite some advances in the attempt to identify how the anatomical differences occur, the restriction of enzymes to specific cells, and how shifts in gene expression have evolved during the transition to C_4 photosynthesis in Cleomaceae (Bayat et al., 2018), little is known about the subject. It is believed that this data limitation is due to the comparison of species that have different evolutionary histories (*e.g. G. gynandra* – Old Word and *T. hassleriana* – New World; differents clusters – Fig. 1). In this regard, it would be more appropriate to compare the species of the same cluster, such as the genus *Coalisina*, which has species with different photosynthetic mechanisms (C_3, C_3-C_4 and C_4). In addition, evaluating several accessions of the same species (*e.g. G. gynandra* - Reeves et al., 2018) can support future genetic mapping to identify causal genes underlying standing natural variation in the componentes of C_4 photosynthesis (Reeves et al., 2018). Thus, based on what was exposed, studies with the family aiming understand the evolution of the C_4 mechanism in Cleomaceae are still scarce, and much of that is known from studies with other plants (*e.g. Flaveria* - Drincovich et al., 1998; Heckmann et al., 2013; McKown et al., 2005; Mckown and Dengler, 2007; Sage et al., 2013 and *Moricandia* - Rawsthorne et al., 1988; Schlüter et al., 2017; Schluter and Weber, 2016).

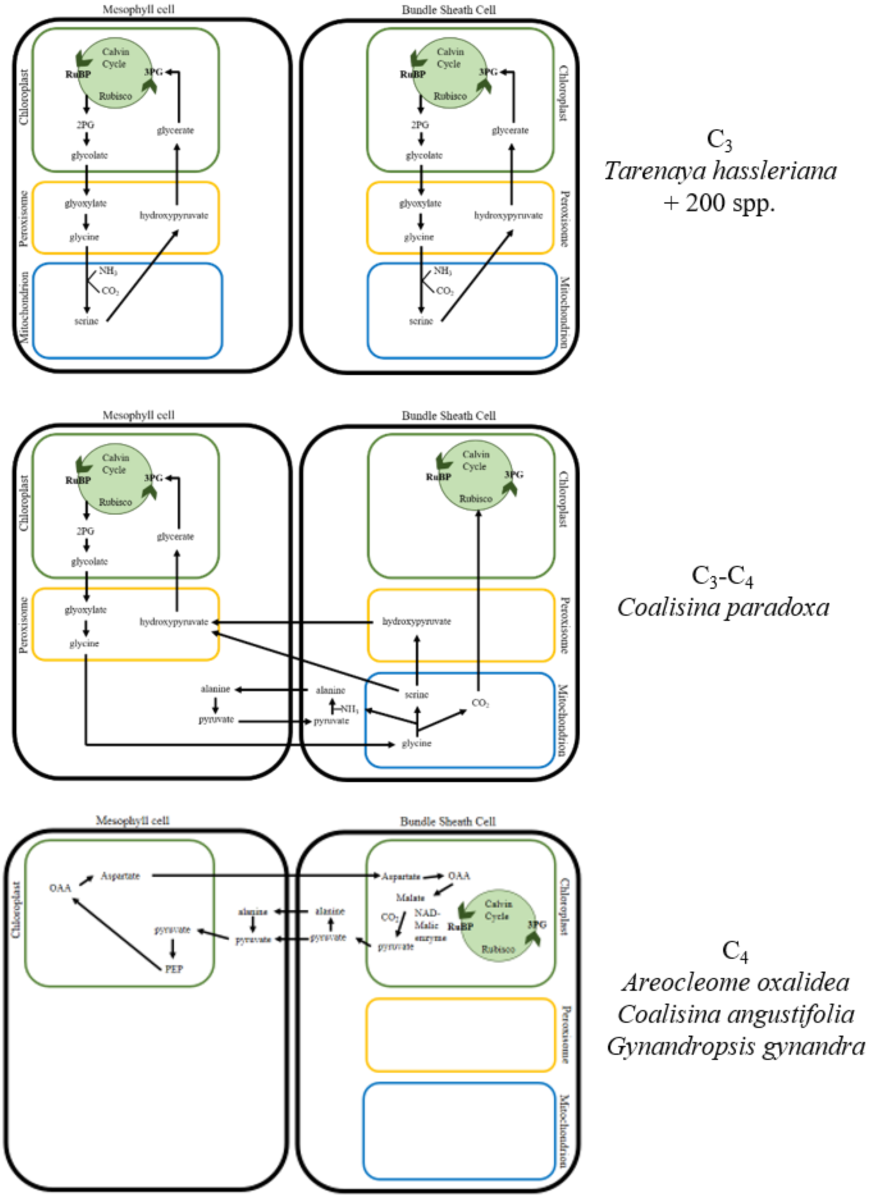


Figure 2: Comparison of schematic pathways of C_3, C_3-C_4 and C_4 photosynthesis in Cleomaceae species. In C_3 species, the Calvin-Belson cycle and photorespiration work independently in mesophyll cell (M) and bundle sheath cells (BSC). In C_3-C_4 species the activity of the glycine decarboxylase complex has shifted towards the BSC, activating CO_2 transport from de M to the BSC by a photorespiratory glycine pump. In C_4 species, the RuBisCO activity shifted completely to the BSC; in Cleomaceae the NAD-ME enzyme is used.

USES, PRODUCTIVITY AND SEED GERMINATION

The Cleomaceae family has some species with socio-economic, medicinal and ornamental importance (Anburaj et al., 2011; Rathore et al., 2013; Singh et al., 2018). Among the species with socioeconomic importance *G. gynandra* stands out, as it plays an important role in the food and nutrition security sub-Saharan Africa and Asia countries and are cultivated on dry areas (Wu et al., 2018), with a potential to improve food security and micronutrient deficiencies (Omondi et al., 2017). Some species, such as *T. aculeta* (Oliveira-Júnior and Conceição, 2010), *T. longicarpa* (Gonc et al., 2006), *Cleome rutidosperma* (Okonwu et al., 2017), *C. viscosa* (Singh et al. 2015) have been used as curative agents often used as conventional drugs to treat several illnesses (Singh et al., 2018). In addition, many species produce essential oils that exhibit antifungal, antimicrobial and insecticidal activity. Thus, the plant extract of some Cleomaceae has inhibitory effects on the growth of different organisms considered as pests in different cultures and are a good resource of natural bioactive antioxidants. Among the species with these properties we can mention, *C. droserifolia*, *C. trinervia* (Muhaidat et al., 2015), *C. rutidosperma* (McNeil et al., 2018), *C. brachycarpa* (Rassouli et al., 2014), *C. amblyocarpa* (Shahin et al., 2018), *C. serrata* (McNeil et al., 2012), *C. heratensis* (Nasseri et al., 2017, 2019), *C. iberica* (Moridi Farimani et al., 2017), *G. gynandra* (Nyalala and Grout, 2007) and *T. spinosa* (Da Silva et al., 2016; McNeil et al., 2010). In turn, *T. hassleriana* and *T. spinosa* (Tsai et al., 2012) are one of the species considered ornamental, widely cultivated in gardens and botanical gardens around the world.

Although the family is recognized as important in several fields, there are few studies on the productivity of this group of plants (Chweya and Mnzava, 1997; Ochuodho et al., 2012; Oluoch et al., 2009; Houdegbe et al., 2018). In this sense, to accelerate the domestication of Cleomaceae species and large-scale production it is necessary to develop adequate agronomic practices for yield improvement (Houdegbe et al., 2018; Onyango et al., 2013). Thus, studies aiming the agronomic related issues such as planting density, fertilization, pest management, irrigation need, transplanting time and cutting modes on the yield are still needed (Houdegbe et al., 2018). Studies in this line have been developed only with *G. gynandra* (Chweya and Mnzava, 1997; Ochuodho et al., 2012; Oluoch et al., 2009; Houdegbe et al., 2018). In this sense, Houdegbe et al. (2018) verified for species a density of ca. 445.000 plants ha^{-1} (spacing of 15 cm×15 cm), and, as well as Chweya and Mnzava, (1997), Ochuodho et al., (2012), Oluoch et al., (2009) observed the average yield around 30 tons ha^{-1}. It should be noted that the production in biomass validated for *G. gynandra* is significant, especially when we consider other crops and even other vegetables, since this vegetable is consumed as such. In this regard, the knowledge of beneficial nutrient contents is an incentive for promoting consumption for improved human health (Omondi et al., 2017).

In addition to studies on productivity, there is also a deficiency in the area of germination. It is known that some species have dormancy and some processes are necessary to increase efficiency in the germination rate. In this sense, for example, for *G. gynandra,*

Table 1: Physiological variation for phtosynthetic gas-exchange and anatomical parameters among diverse species of Cleomaceae.

Species*	**CO_2 Γ** (µmol mol^{-1})	A_N (µmol m^{-2}s^{-1})	**δ^{13}C (‰)**	**Vein density (mm mm^{-2})**	**Reference**
Areocleome oxalídea - C_4	-	-	-14.1 - 15.8	-	(Koteyeva et al., 2011; Marshall et al., 2007; Voznesenskaya et al., 2007)
Cleome africana - C_3	52.4 – 65.0	-	-30.1	10	(Koteyeva et al., 2011; Marshall et al., 2007; Voznesenskaya et al., 2007)
Peritoma arborea - C_3	49.0	-	-25.2 - 27.5	7	(Marshall et al., 2007; Voznesenskaya et al., 2007)
*Cleome marshallii*** - C_3	52.0	-		7	(Marshall et al., 2007)
Cleome ornithodioides - C_3	66.5	-	-28.0 - 29.7	-	(Voznesenskaya et al., 2007)
Cleome violacea - C_3	51.0	-	-29.5	8	(Marshall et al., 2007; Voznesenskaya et al., 2007)
Coalisina angustifolia - C_4	-	-	-12.5 - 13.2	-	(Koteyeva et al., 2011; Marshall et al., 2007; Voznesenskaya et al., 2007)
Coalisina paradoxa - C_3-C_4	27.5	20	-23.8 – 26.0	11	(Koteyeva et al., 2011; Marshall et al., 2007; Voznesenskaya et al., 2007)
Gynandropsis gynandra - C_4	2.0-9.0	23-35	-13.9 – 17.1	6-10	(Koteyeva et al., 2011; Marshall et al., 2007; Reeves et al., 2018; Voznesenskaya et al., 2007)
Kersia foliosa - C_3	40.0	-	-22.8 - 25.6	8	(Marshall et al., 2007; Voznesenskaya et al., 2007)
Melidiscus giganteus - C_3	66.9	-	-23.4 - 29.3	-	(Voznesenskaya et al., 2007)
Sieruela hirta - C_3	123.0	-	-25.7 - 28.8	7	(Marshall et al., 2007; Voznesenskaya et al., 2007)
Sieruela monophylla - C_3	62.0	23	-27.8	-	(Voznesenskaya et al., 2007)
Tarenaya hassleriana - C_3	62.5	-	-24.6 - 29.8	-	(Marshall et al., 2007; Voznesenskaya et al., 2007)
Tarenaya spinosa - C_3	50.0-66.7	-	-26.41 - 30.6	5	(Marshall et al., 2007; Voznesenskaya et al., 2007)

*All names have been updated according to IPNI 2020. ** invalid name.

Zohoungbogbo et al., (2018) observed only 25% and Ekpong, (2009) 17% in the germination rate of the seeds without any previous treatment. Other species such as *T. longicarpa* (De Castro et al., 2014), *T. hassleriana* and *Polanisia dodecandra* (Gomez Raboteaux and Anderson, 2010), *Cleomella lutea* and *C. serrulata* (Cane, 2008) seed dormancy was also observed. However, these studies increased the germination rates of species (~90%) by leaving a longer storage time for the seeds, the use of alternative temperatures, as well different treatments including various levels of GA3, KNO3, leaching, pre-chilling, soaking and pre-heating at different temperatures (Ekpong, 2009).

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Cleomaceae is a promising family in the answers that researchers have long sought. The family, in this regard, can include studies in several areas such as the evolution of C_4 photosynthesis. In this sense, to increase the productivity of C_3 species under conditions in which their development is not favorable. Furthermore, In addition, the family can assist in the production of production of hybrids that can attract more pollinators and conserve them or even be spontaneously self-pollinated. In addition, the family have the potential in the development of drugs for various diseases and can be used even as insecticides, besides some species can be consumed as vegetables due to its great nutritional potential. Thus, there are many opportunities to advance studies in different areas with the family; however, basic studies such as describing the existing diversity and understanding the existing genetics are still scarce. These studies are important for the preservation/protection of species, establishment of protocols for in vitro cultivation and transformation, creation of a germplasm bank, among others.

ACKNOWLEDGMENT

This work was supported by funding from the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) (grant number 424024/2018-7), the Foundation for Research Assistance of the Minas Gerais State (Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG) (grant number APQ-00528-18 and CRA-RED00053-16). Research fellowships granted by CNPq to ANN and WLA as well as the scholarship granted by the Coordination of Improvement of Higher Education Personnel (CAPES) to DFP are also gratefully acknowledged. MGMVV was supported by scholarships from CAPES (PNPD-1638006) and FAPEMIG (BPD-00514-14).

APMW acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2048/1 – project ID 390686111.

REFERENCES

Aizen, M.A., Garibaldi, L.A., Cunningham, S.A., Klein, A.M., 2009. How much does agriculture depend on pollinators? Lessons from long-term trends in crop production. Ann. Bot. 103, 1579–1588. https://doi.org/10.1093/aob/mcp076

Al-Shehbaz, I. A. 1973. The biosystematics of the genus Thelypodium (Cruciferae). Contr. Gray Herb. 204: 3-148.

Ammal, J.E.K., 1933. The chromosome number of Cleome viscosa Linn. Curr. Sci. 328–328. https://doi.org/10.1016/B978-0-12-409548-9.10607-4

Anburaj, J., Singh, C.R., Sundarraj, S., Kannan, S., 2011. In vitro regeneration of Cleome viscosa – an important medicinal herb 9, 37–44.

Barrett, R.L., Roalson, E.H., Ottewell, K., Byrne, M., Govindwar, S.P., Yadav, S.R., Tamboli, A.S., Gholave, A.R., 2017. Resolving Generic Boundaries in Indian-Australasian Cleomaceae: Circumscription of *Areocleome*, *Arivela*, and *Corynandra* as Distinct Genera. Syst. Bot. 42, 694–708. https://doi.org/10.1600/036364417X696401

Barrett, S.C.H., 2002. The evolution of plant sexual diversity. Nat. Rev. Genet. https://doi.org/10.1038/nrg776

Bayat, S., Schranz, M.E., Roalson, E.H., Hall, J.C., 2018b. Lessons from Cleomaceae, the Sister of Crucifers. Trends Plant Sci. 23, 808–821. https://doi.org/10.1016/j.tplants.2018.06.010

Bergh, E. Van Den, Külahoglu, C., Bräutigam, A., Hibberd, J.M., Weber, A.P.M., Zhu, X., Schranz, M.E., 2014. Current Plant Biology Gene and genome duplications and the origin of C 4 photosynthesis : Birth of a trait in the Cleomaceae. Curr. Plant Biol. 1, 2–9. https://doi.org/10.1016/j.cpb.2014.08.001

Bradley, D., Carpenter, R., Sommer, H., Hartley, N., Coen, E., 1993. Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the plena locus of antirrhinum. Cell 72, 85–95. https://doi.org/10.1016/0092-8674(93)90052-R

Brown, N.J., Parsley, K., Hibberd, J.M., 2005. The future of C4 research - Maize, Flaveria or Cleome? Trends Plant Sci. 10, 215–221. https://doi.org/10.1016/j.tplants.2005.03.003

Bruijn, S. De, Bhide, A.S., Kuelahoglu, C., Bian, C., Chen, J., Fan, G., Kaufmann, K., Hall, J.C., Becker, A., Bräutigam, A., Weber, A.P.M., Shi, C., 2013. The Tarenaya hassleriana Genome Provides Insight into Reproductive Trait and Genome Evolution of Crucifers 25, 2813–2830. https://doi.org/10.1105/tpc.113.113480

Cane, J.H., 2008. Breeding biologies , seed production and species-rich bee guilds of Cleome lutea and Cleome serrulata (Cleomaceae) 152–158. https://doi.org/10.1111/j.1442-1984.2008.00224.x

Causier, B., Schwarz-Sommer, Z., Davies, B., 2010. Floral organ identity: 20 years of ABCs. Semin. Cell Dev. Biol. https://doi.org/10.1016/j.semcdb.2009.10.005

Cheng, S., van den Bergh, E., Zeng, P., Zhong, X., Xu, J., Liu, X., Hofberger, J., de Bruijn, S., Bhide, A.S., Kuelahoglu, C., Bian, C., Chen, J., Fan, G., Kaufmann, K., Hall, J.C., Becker, A., Bräutigam, A., Weber, A.P.M., Shi, C., Zheng, Z., Li, W., Lv, M., Tao, Y., Wang, J., Zou, H., Quan, Z., Hibber, J.M., Zhang, G., Zhu, X.G., Xu, X., Schranz, M.E., 2013. The Tarenaya hassleriana genome provides insight into reproductive trait and genome evolution of crucifers. Plant Cell 25, 2813–2830. https://doi.org/10.1105/tpc.113.113480

Christin, P.A., Besnard, G., Edwards, E.J., Salamin, N., 2012. Effect of genetic convergence on phylogenetic inference. Mol. Phylogenet. Evol. 62, 921–927. https://doi.org/10.1016/j.ympev.2011.12.002

Christin, P.A., Osborne, C.P., Chatelet, D.S., Columbus, J.T., Besnard, G., Hodkinson, T.R., Garrison, L.M., Vorontsova, M.S., Edwards, E.J., 2013. Anatomical enablers and the evolution of C4 photosynthesis in grasses. Proc. Natl. Acad. Sci. U. S. A. 110, 1381–1386. https://doi.org/10.1073/pnas.1216777110

Christin, P.A., Sage, T.L., Edwards, E.J., Ogburn, R.M., Khoshravesh, R., Sage, R.F., 2011. Complex evolutionary transitions and the significance of c3-c4 intermediate forms of photosynthesis in molluginaceae. Evolution (N. Y). 65, 643–660. https://doi.org/10.1111/j.1558-5646.2010.01168.x

Chweya, J.A., Mnzava, N.A., 1997. Cat's whiskers. Cloeme gynandra L. Promoting the Conservation and Use of Underutilized and Neglected Crops. 11. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome, Italy.

Costa e Silva, M. B. 2000. O Genero Cleome L. (Capparaceae Juss.) para o Brasil. Unpublished Ph.D. Dissertation, Universidade Federal Rural de Pernambuco, Recife, Brazil.

Da Silva, A.P.S.A., Nascimento Da Silva, L.C., Martins Da Fonseca, C.S., de Araújo, J.M., Correia, M.T. do. S., Cavalcanti, M. da S., Lima, V.L. d. M., 2016. Antimicrobial activity and phytochemical analysis of organic extracts from Cleome spinosa Jaqc. Front. Microbiol. 7, 963. https://doi.org/10.3389/fmicb.2016.00963

Davies, B., Motte, P., Keck, E., Saedler, H., Sommer, H., Schwarz-Sommer, Z., 1999. PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. EMBO J. 18, 4023–34. https://doi.org/10.1093/emboj/18.14.4023

De Candolle AP (1824) Prodromus Systematis Naturalis Regni Vegetabilis 1:241–242.

Ducke A (1930) Plantes nouvelles ou peu connues de la region amazonienne. Archiv Jard Bot Rio de Janeiro 5:115–116.

De Bruijn, S., Zhao, T., Muiño, J.M., Schranz, E.M., Angenent, G.C., Kaufmann, K., 2018. PISTILLATA paralogs in Tarenaya hassleriana have diverged in interaction specificity. BMC Plant Biol. 18, 1–18. https://doi.org/10.1186/s12870-018-1574-0

De Castro, T.C., Simões-Gurgel, C., Ribeiro, I.G., Coelho, M.G.P., Albarello, N., 2014. Aspectos morfológicos de frutos, sementes, plântulas e germinação in vivo e in vitro de espécies do gênero cleome. J. Seed Sci. 36, 326–335. https://doi.org/10.1590/2317-1545v36n31013

Drincovich, M.F., Casati, P., Andreo, C.S., Chessin, S.J., Franceschi, V.R., Edwards, G.E., Ku, M.S.B., 1998. Evolution of C4 photosynthesis in flaveria species: Isoforms of NADP-malic enzyme. Plant Physiol. 117, 732–744. https://doi.org/10.1104/pp.117.3.733

Eichler A (1865) Capparidaceae. In: Martius (ed) Flora Brasiliensis 13(1):238–292.

Ekpong, B., 2009. Scientia Horticulturae Effects of seed maturity , seed storage and pre-germination treatments on seed germination of cleome (Cleome gynandra L .) 119, 236–240. https://doi.org/10.1016/j.scienta.2008.08.003

Feodorova, Tatiana A., Voznesenskaya, E. V., Edwards, G.E., Roalson, E.H., 2010. Biogeographic Patterns of Diversification and the Origins of Cleome (Cleomaceae). Syst. Bot. 35, 811–826. https://doi.org/10.1600/036364410X539880

Feodorova, Tatiana A, Voznesenskaya, E. V, Edwards, G.E., Roalson, E.H., Feodorova, T.A., Voznesenskaya, E. V, Edwards, G.E., Roalson, E.H., 2010. Biogeographic Patterns of Diversification and the Origins of C_4 in Cleome (Cleomaceae). 35, 811–826. https://doi.org/10.1600/036364410X539880

Gomez Raboteaux, N.N., Anderson, N.O., 2010. Germination of Cleome hassleriana and Polanisia dodecandra seed lots in response to light, temperature and stratification. Res. J. Seed Sci. 3, 1–17. https://doi.org/10.3923/rjss.2010.1.17

Gonc, P.F., Albarello, N., Simo, C., Rosas, A., Castro, T.C.D.E., 2006. In vitro propagation of *Cleome spinosa* (Capparaceae) using explants from nursery-grown seedlings and axenic plants 601–606. https://doi.org/10.1079/IVP2006828

Gowik, U., Westhoff, P., 2011. The Path from C_3 to C_4 Photosynthesis 1 155, 56–63. https://doi.org/10.1104/pp.110.165308

Hall, J.C., 2008. Systematics of Capparaceae and Cleomaceae : an evaluation of the generic delimitations of Capparis and Cleome using plastid DNA sequence data 1 696, 682–696. https://doi.org/10.1139/B08-026

Hall, Jocelyn C., Sytsma, K.J., Iltis, H.H., 2002. Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. Am. J. Bot. 89, 1826–1842. https://doi.org/10.3732/ajb.89.11.1826

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., de Kroon, H., 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLoS One 12, e0185809. https://doi.org/10.1371/journal.pone.0185809

Heilborn O (1931) Section Fruticosae Eichl. of the genus Cleome L. Arkiv For Botanik 23(10):1–19.

Hutchinson J (1967) The genera of flowering plants 2. (Capparidaceae). Clarendon Press, Oxford, pp 303–317.

Heckmann, D., Schulze, S., Denton, A., Gowik, U., Westhoff, P., Weber, A.P.M., Lercher, M.J., 2013. Theory Predicting C 4 Photosynthesis Evolution : Modular , Individually Adaptive Steps on a Mount Fuji Fitness Landscape. Cell 153, 1579–1588. https://doi.org/10.1016/j.cell.2013.04.058

Houdegbe, C.A., Sogbohossou, E.O.D., Achigan-Dako, E.G., 2018. Enhancing growth and leaf yield in Gynandropsis gynandra (L.) Briq. (Cleomaceae) using agronomic practices to accelerate crop domestication. Sci. Hortic. (Amsterdam). 233, 90–98. https://doi.org/10.1016/j.scienta.2018.01.035

Iltis HH (1952) A revision of the New World species of Cleome. Ph.D. thesis, Washington University, St Louis, 335 p

Iltis, H.H., Hall, J.C., Cochrane, T.S., Sytsma, K.J., 2011. Studies in the Cleomaceae I. On the Separate Recognition of Capparaceae, Cleomaceae, and Brassicaceae [1]. Ann. Missouri Bot. Gard. 98, 28–36. https://doi.org/10.3417/2007017

Imbamba, S.K., Tieszen, L.L., 1977. Influence of Light and Temperature on Photosynthesis and Transpiration in Some C3 and C4 Vegetable Plants from Kenya. Physiol. Plant. 39, 311–316. https://doi.org/10.1111/j.1399-3054.1977.tb01890.x

Inda, L.A., Torrecilla, Æ.P., Catala, Æ.P., Ruiz-zapata, T., 2008. Phylogeny of Cleome L. and its close relatives Podandrogyne Ducke and Polanisia Raf . (Cleomoideae , Cleomaceae) based on analysis of nuclear ITS sequences and morphology 111–126. https://doi.org/10.1007/s00606-008-0026-y

Jack, T., Brockman, L.L., Meyerowitz, E.M., 1992. The Homeotic Gene APETALA3 of Arabidopsis thaliana Encodes a MADS Box and Is Expressed in Petals and Stamens 68, 683–697.

Keller S (1979) A revision of the genus Wislizenia (Capparaceae) based on populations studies. Brittonia 31:333–351.

Kers LE (2003) Capparaceae. In: Kubitzki K, Bayer C (eds) Flowering plants. Dicotyledons. Malvales, Capparales and Non-betalain Caryophyllales. In: Kubitzki K (ed) The families and genera of vascular plants, vol 5. Springer, Heidelberg, pp 36–56.

Koteyeva, N.K., Voznesenskaya, E. V, Roalson, E.H., Edwards, G.E., 2011. Diversity in forms of C 4 in the genus Cleome (Cleomaceae) 269–283. https://doi.org/10.1093/aob/mcq239

Krenzer, E.G., Moss, D.N., Crookston, R.K., 1975. Carbon Dioxide Compensation Points of Flowering Plants. Plant Physiol. 56, 194–206. https://doi.org/10.1104/pp.56.2.194

Külahoglu, C., Denton, A.K., Sommer, M., Maß, J., Schliesky, S., Wrobel, T.J., Berckmans, B., Gongora-castillo, E., Buell, C.R., Simon, R., Veylder, L. De, Bräutigam, A., Weber, A.P.M., 2014. Comparative Transcriptome Atlases Reveal Altered Gene Expression Modules between Two Cleomaceae C3 and C4 Plant Species 26, 3243–3260. https://doi.org/10.1105/tpc.114.123752

Levin, D.A., 2002. The role of chromosomal change in plant evolution. Oxford University Press.

Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowmant, J.L., Yanofsky, M.F., 2000. SHATTERPROOF MADS-box genes control dispersal in Arabidopsis. Nature 404, 766–770. https://doi.org/10.1038/35008089

Luciano, R., Neto, S., Thomas, W.W., Regina, M., Barbosa, D.V., Roalson, E.H., 2018. New combinations and taxonomic notes for Tarenaya (Cleomaceae) 32, 540–545. https://doi.org/10.1590/0102-33062017abb0417

Lundgren, M.R., Osborne, C.P., Christin, P.A., 2014. Deconstructing Kranz Anatomy to Understand C4 Evolution - PubMed. J. Exp. Bot. 65, 3357–3369.

Machado, I., Cristina Lopes, A., Valentina Leite, A., Virgíniade Brito Neves, C., 2006. Cleome spinosa (Capparaceae): polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. Bot. Jahrbücher für Syst. Pflanzengeschichte und Pflanzengeographie 69–82. https://doi.org/10.1127/0006-8152/2006/0127-0069

Marshall, D.M., Muhaidat, R., Brown, N.J., Liu, Z., Stanley, S., Griffiths, H., Sage, R.F., Hibberd, J.M., 2007. Cleome, a genus closely related to Arabidopsis , contains species spanning a developmental progression from C 3 to C 4 photosynthesis 886–896. https://doi.org/10.1111/j.1365-313X.2007.03188.x

Marshman, J., Blay-Palmer, A., Landman, K., 2019. Anthropocene Crisis: Climate Change, Pollinators, and Food Security. Environments 6, 22. https://doi.org/10.3390/environments6020022

Mckown, A.D., Dengler, N.G., 2007. Key innovations in the evolution of Kranz anatomy and C4 vein pattern in Flaveria (Asteraceae). Am. J. Bot. 94, 382–389. https://doi.org/10.3732/ajb.94.3.382

McKown, A.D., Moncalvo, J.M., Dengler, N.G., 2005. Phylogeny of Flaveria (Asteraceae) and inference of C4 photosynthesis evolution. Am. J. Bot. 92, 1911–1928. https://doi.org/10.3732/ajb.92.11.1911

McNeil, M.J., Porter, R.B.R., Rainford, L., Dunbar, O., Francis, S., Laurieri, N., Delgoda, R., 2018. Chemical composition and biological activities of the essential oil from Cleome rutidosperma DC. Fitoterapia 129, 191–197. https://doi.org/10.1016/j.fitote.2018.07.006

McNeil, M.J., Porter, R.B.R., Williams, L.A.D., Rainford, L., 2010. Chemical composition and antimicrobial activity of the essential oils from Cleome spinosa. Nat. Prod. Commun. 5, 1301–1306. https://doi.org/10.1177/1934578x1000500833

Moridi Farimani, M., Nazarianpoor, E., Rustaie, A., Akhbari, M., 2017. Phytochemical constituents and biological activities of Cleome iberica DC. Nat. Prod. Res. 31, 1329–1332. https://doi.org/10.1080/14786419.2016.1239093

Nasseri, M., Behravesh, S., Allahresani, A., 2017. Essential oil composition of Cleome heratensis (Capparaceae) at different growing stages. Q. J. Iran. Chem. Commun.

Nasseri, M.A., Behravesh, S., Allahresani, A., Kazemnejadi, M., 2019. Phytochemical and antioxidant studies of Cleome heratensis (Capparaceae) plant extracts. Bioresour. Bioprocess. 6, 1–10. https://doi.org/10.1186/s40643-019-0240-1

Novais, S.M.A., Nunes, C.A., Santos, N.B., D'Amico, A.R., Fernandes, G.W., Quesada, M., Braga, R.F., Neves, A.C.O., 2016. Effects of a possible pollinator crisis on food crop production in Brazil. PLoS One 11. https://doi.org/10.1371/journal.pone.0167292

Nozzolillo, C., Treyvaud, V., Bily, A.C., Harris, C.S., Saleem, A., Andersen, Ø.M., Jordheim, M., 2010. Novel aspects of the fl owers and fl oral pigmentation of two Cleome species (Cleomaceae), C . hassleriana and C . serrulata 38, 361–369. https://doi.org/10.1016/j.bse.2010.03.005

Nyalala, S., Grout, B., 2007. African spider flower (Cleome gynandra L./Gynandropsis gynandra (L .) Briq .) as a red spider mite (Tetranychus urticae Koch) repellent in cut-flower rose (Rosa hybrida L.) cultivation 114, 194–198. https://doi.org/10.1016/j.scienta.2007.06.010

Okonwu, K., Ekeke, C., Mensah, S.I., 2017. Micromorphological and Phytochemical Studies on Cleome rutidosperma Linn . 11, 1–8. https://doi.org/10.9734/JABB/2017/31028

Ochuodho, J.O., K'Opondo, F.B.O., Ojiewo, C.O., 2012. Good agricultural practices for production of spiderplant in sub-Saharan Africa. In: Nono-Womdim, R., Ojiewo, C., Abang, M., Oluoch, M.O. (Eds.), Good Agricultural Practices for African Indigenous Vegetables. ISHS, Gent-Oostakker, Belgium, pp. 75–82.

Oluoch, M.O., Pichop, G.N., Drissa, S., Abukutsa-Onyango, M.O., Diouf, M., Shackleton, C.M., 2009. Production and harvesting systems for african indigenous vegetables. In: Shackleton, C.M., Pasquini, M.W., Drescher, A.W. (Eds.), African Indigenous Vegetables in Urban Agriculture. Earthscan, UK and USA, pp. 145–176.

Oliveira-Júnior, S.R., Conceição, G.M., 2010. Espécies vegetais nativas do cerrado utilizadas como medicinais pela comunidade brejinho, Caxias, Maranhão, Brasil. Cadernos de Geociências. Cad. Geociências 7, 140–148.

Omondi, E.O., Debener, T., Linde, M., Abukutsa-Onyango, M., Dinssa, F.F., Winkelmann, T., 2017a. Mating biology, nuclear DNA content and genetic diversity in spider plant (*Cleome gynandra*) germplasm from various African countries. Plant Breed. 136, 578–589. https://doi.org/10.1111/pbr.12485

Omondi, E.O., Engels, C., Nambafu, G., Schreiner, M., Neugart, S., Abukutsa-Onyango, M., Winkelmann, T., 2017b. Nutritional compound analysis and morphological characterization of spider plant (Cleome gynandra) - an African indigenous leafy vegetable. Food Res. Int. 100, 284–295. https://doi.org/10.1016/j.foodres.2017.06.050

Onyango, C.M., Kunyanga, C.N., Ontita, E.G., Narla, R.D., Kimenju, J.W., 2013. Current status on production and utilization of spider plant (Cleome gynandra L.) an underutilized leafy vegetable in Kenya. Genet. Resour. Crop Evol. 60, 2183–2189. https://doi.org/10.1007/s10722-013-0036-7

Patchell, M.J., Bolton, M.C., Mankowski, P., Hall, J.C., 2011. Comparative Floral

Development in Cleomaceae Reveals Two Distinct Pathways Leading to Monosymmetry. Int. J. Plant Sci. 172, 352–365. https://doi.org/10.1086/658158

Patchell, M.J., Roalson, E.H., Hall, J.C., 2014. Resolved phylogeny of Cleomaceae based on all three genomes 63, 315–328.

Pax F, Hoffmann K (1936) Capparidaceae. In: Engler A, Prantl K. (eds) Die Natu¨rlichen Pflanzenfamilien. 2nd edn. 17 b: Leipzig, pp 146–223.

Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E., Yanofsky, M.F., 2000. B and C floral organ identity functions require SEPALLATTA MADS-box genes. Nature 405, 200–203. https://doi.org/10.1038/35012103

Prasifka, J.R., Mallinger, R.E., Portlas, Z.M., Hulke, B.S., Fugate, K.K., Paradis, T., Hampton, M.E., Carter, C.J., Campbell, A.J., 2018. Using Nectar-Related Traits to Enhance Crop-Pollinator Interactions 9, 1–8. https://doi.org/10.3389/fpls.2018.00812

Raghavendra, A.S., Das, V.S.R., 1978. Occurrence of C4 photosynthesis: a supplementary list of C4 plants reported during late 1974 mid 1977. Photosynthetica 12, 200–208.

Rajendrudu, G., Rama Das, V.S., 1982. The carboxylating enzymes in leaves of Cleome gynandra, A C4 dicot plant. Plant Sci. Lett. 26, 285–291. https://doi.org/10.1016/0304-4211(82)90103-1

Rassouli, E., Dadras, O.G., Bina, E., Asgarpanah, J., 2014. The Essential Oil Composition of Cleome brachycarpa Vahl ex DC. J. Essent. Oil-Bearing Plants 17, 158–163. https://doi.org/10.1080/0972060X.2014.884784

Rathore, N.S., Rathore, N., Shekhawat, N.S., 2013. In vitro flowering and seed production in regenerated shoots of Cleome viscosa. Ind. Crop. Prod. 50, 232–236. https://doi.org/10.1016/j.indcrop.2013.07.032

Rawsthorne, S., Hylton, C.M., Smith, A.M., Woolhouse, H.W., 1988. Distribution of photorespiratory enzymes between bundle-sheath and mesophyll cells in leaves of the C3-C4 intermediate species Moricandia arvensis (L.) DC. Planta 176, 527–532. https://doi.org/10.1007/BF00397660

Reeves, G., Singh, P., Rossberg, T.A., Sogbohossou, E.O.D., Schranz, M.E., Hibberd, J.M., 2018. Natural variation within a species for traits underpinning C 4 photosynthesis. Plant Physiol. 177, 504–512. https://doi.org/10.1104/pp.18.00168

Reyna-llorens, I., Hibberd, J.M., Hibberd, J.M., 2017. Recruitment of pre-existing networks during the evolution of C 4 photosynthesis 2–7.

Riaz, S., Abid, R., Syed, A., Ali, S.A., Munir, I., Qaiser, M., 2019. Morphology and seed protein profile for a new species of the genus Cleome L. (Cleomaceae) from Pakistan. Protein Chem. to Proteomics 1–10.

Roalson, E.H., Hall, J.C., Ii, J.P.R., Cardinal-mcteague, W.M., Cochrane, T.S., Sytsma, K.J., 2015. A revision of generic boundaries and nomenclature in the North American cleomoid clade (Cleomaceae) 205, 2–6.

Roalson, E.H., Hall, J.C., Roalson, E.H., Hall, J.C., n.d. New Generic Concepts for African Cleomaceae New Generic Concepts for African Cleomaceae 42, 925–942. https://doi.org/10.1600/036364417X696393

Ruiz Zapata T. & N. Xena de Enrech. 1997. La morfologia del polen de Cleome L. (Capparidaceae) en relacion con su taxonomya y sindromes de polinizacion. Acta Bot. Venez. 20: 67-80.

Sage, R.F., 2004. The evolution of C4 photosynthesis. New Phytol. https://doi.org/10.1111/j.1469-8137.2004.00974.x

Sage, R.F., 2001. Environmental and evolutionary preconditions for the origin and diversification of the C4 photosynthetic syndrome. Plant Biol. https://doi.org/10.1055/s-2001-15206

Sage, T.L., Busch, F.A., Johnson, D.C., Friesen, P.C., Stinson, C.R., Stata, M., Sultmanis, S., Rahman, B.A., Rawsthorne, S., Sage, R.F., 2013. Initial Events during the Evolution of C4 Photosynthesis in C3 Species of Flaveria. Plant Physiol. 163.

Sanchez-Acebo, L., 2005. A Phylogenetic Study of the New World Cleome (Brassicaceae, Cleomoideae) on JSTOR. Ann. Missouri Bot. Gard. 92, 179–201.

Sankhla, N., Ziegler, H., Vyas, O.P., Stichler, W., Trimborn, P., 1975. Eco-physiological studies on Indian arid zone plants - V. A. Screening of some species for the C4-pathway of photosynthetic CO2-fixation. Oecologia 21, 123–129. https://doi.org/10.1007/BF00345555

Schlüter, U., Bräutigam, A., Gowik, U., Melzer, M., Christin, P., Kurz, S., Mettler-altmann, T., Weber, A.P.M., 2017. Photosynthesis in C3 – C4 intermediate Moricandia species 68, 191–206. https://doi.org/10.1093/jxb/erw391

Schranz, M.E., Mitchell-olds, T., 2006. Independent Ancient Polyploidy Events in the Sister Families Brassicaceae and Cleomaceae 18, 1152–1165. https://doi.org/10.1105/tpc.106.041111.1

Shahin, S., Kurup, S., Cheruth, A.J., Salem, M., 2018. Chemical Composition of Cleome amblyocarpa Barr. & Murb. Essential Oils Under Different Irrigation Levels in Sandy

Soils with Antioxidant Activity. J. Essent. Oil-Bearing Plants 21, 1235–1256. https://doi.org/10.1080/0972060X.2018.1512422

Singh, H., Mishra, A., Mishra, A.K., 2018b. The chemistry and pharmacology of Cleome genus: A review. Biomed. Pharmacother. 101, 37–48. https://doi.org/10.1016/j.biopha.2018.02.053

Soares-Neto, L., Thomas, W.W., Barbosa, M.R.V., Roalson, E.H. 2020a. Diversification of new world Cleomaceae with emphasis on *Tarenaya* and the description of a new genus.

Soares-Neto, L., Thomas, W.W., Barbosa, M.R.V., Roalson, E.H. 2020b. A well-known "mussambê" is a new species of Tarenaya (Cleomaceae) from South America.

Stevens, P.F., n.d. Angiosperm Phylogeny Website. Version 14, July 2017 [and more or less continuously updated since] [WWW Document]. URL http://www.mobot.org/mobot/research/apweb/orders/brassicalesweb.htm (accessed 2.27.19).

Subramania, D., Susheela, G., n.d. Cytotaxonomical Studies of South Indian Capparidaceae.

Thulin, M., Roalson, E.H., Thulin, M., Roalson, E.H., n.d. Resurrection of the Genus Rorida (Cleomaceae), a Distinctive Old World Segregate of Cleome Resurrection of the Genus Rorida (Cleomaceae), a Distinctive Old World Segregate of Cleome 42, 569–577. https://doi.org/10.1600/036364417X695989

Tsai, Y.T., Chen, P.Y., To, K.Y., 2012. Plant regeneration and stable transformation in the floricultural plant Cleome spinosa, a C 3 plant closely related to the C 4 plant C. gynandra. Plant Cell Rep. 31, 1189–1198. https://doi.org/10.1007/s00299-012-1240-1

U, S., AP, W., 2016. The Road to C4 Photosynthesis: Evolution of a Complex Trait via Intermediary States. Plant Cell Physiol. 57. https://doi.org/10.1093/PCP/PCW009

Voznesenskaya, E. V., Koteyeva, N.K., Chuong, S.D.X., Ivanova, A.N., Barroca, J., Craven, L.A., Edwards, G.E., 2007. Physiological, anatomical and biochemical characterisation of photosynthetic types in genus Cleome (Cleomaceae). Funct. Plant Biol. 34, 247. https://doi.org/10.1071/FP06287

Williams, B.P., Burgess, S.J., Reyna-llorens, I., Knerova, J., Aubry, S., Stanley, S., Hibberd, J.M., 2016. An Untranslated cis -Element Regulates the Accumulation of Multiple C 4 Enzymes in Gynandropsis gynandra 28, 454–465. https://doi.org/10.1105/tpc.15.00570

Wu, T. hor, Solberg, S.O., Yndgaard, F., Chou, Y.Y., 2018. Morphological patterns in a world collection of Cleome gynandra. Genet. Resour. Crop Evol. 65, 271–283. https://doi.org/10.1007/s10722-017-0529-x

Zohoungbogbo, H.P.F., Houdegbe, C.A., Sogbohossou, D.E.O., Tossou, M.G., Maundu, P., Schranz, E.M., Van Deynze, A., Zoundjihekpon, J., Achigan-Dako, E.G., 2018. Andromonoecy in Gynandropsis gynandra (L.) Briq. (Cleomaceae) and effects on fruit and seed production. Genet. Resour. Crop Evol. 65, 2231–2239. https://doi.org/10.1007/s10722-018-0687-5

GENERAL CONCLUSIONS

Cleomaceae is a highly diverse family in which we can observe various types of photosynthetic metabolism. In addition their members also show great morphological differences. However these traits still little explored. In the present study, based on a small sample of the group (21 morphospecies) it was possible to recognize a new taxon within *Tarenaya* and also to separate the *Tarenaya* series considering both morphological and molecular data. Taken together, these data leads us to believe that in the a near future they can also be recognized as distinct taxa (Chapter 1 and 2). Furthermore, we observed that the sampled species have different types of sexual expression, being essentially characterized in four types: andromonoecy, hermaphrodite plants, polygamous and gino-polygamodioic. Evolutionary features, such as hercogamy, dichogamy and unisexual flowers, to avoid self-pollinated, were present in some of them (Chapter 5). Furthermore, it should be noted that the floral diversity and production of rewards observed are important characteristics at times of pollinators population decline.

In the present study, we also demonstrated the potential of Cleomaceae species by offering novel insights into the evolution of C_4 photosynthetic mechanism, particularly the initial transition from a C_3 character state, in *Tarenaya* species. The C_3 species described here exhibit anatomical shifts, such as increases in venation density and enlargement of bundle sheath cells, which may predispose these lines to the development of C_4 photosynthesis (Chapters 3 and 4). Thus, Cleomaceae, for having all this diversity and being a sister family of Brassicaceae, can be helpful for scientific advancement in several areas and, mainly, help to understand how plants adapt quickly to environmental factors and how this may involve in agriculture. Therefore, it may be easier to understand in this family how the C_4 photosynthetic mechanism operates. Further studies could check the feasibility of those factors regarding plant breeding programs.