

Review

Lessons from Cleomaceae, the Sister of Crucifers

Soheila Bayat,^{1,2,3} M. Eric Schranz,⁴ Eric H. Roalson,⁵ and Jocelyn C. Hall^{1,*}

Cleomaceae is a diverse group well-suited to addressing fundamental genomic and evolutionary questions as the sister group to Brassicaceae, facilitating transfer of knowledge from the model *Arabidopsis thaliana*. Phylogenetic and taxonomic revisions provide a framework for examining the evolution of substantive morphological and physiology diversity in Cleomaceae, but not necessarily in Brassicaceae. The investigation of both nested and contrasting whole-genome duplications (WGDs) between Cleomaceae and Brassicaceae allows comparisons of independently duplicated genes and investigation of whether they may be drivers of the observed innovations. Further, a wealth of outstanding genetic research has provided insight into how the important alternative carbon fixation pathway, C₄ photosynthesis, has evolved via differential expression of a suite of genes, of which the underlying mechanisms are being elucidated.

Cleomaceae as an Emerging Model to Its Sister Family Brassicaceae

The plant family Cleomaceae presents a fascinating juxtaposition of diversity compared to its larger and better-known sister family Brassicaceae. Both families have a cosmopolitan distribution and share the same basic floral ground plan and fruit type [1], but the differences between these families are striking and provide grounds for a range of investigations. The 3700 species of Brassicaceae are mostly distributed in cooler temperate climates, whereas the 270 species of Cleomaceae are more frequently distributed in warm temperate, desert, and tropical areas [2–4]. As the home to model species *A. thaliana* (arabidopsis), many economically important crops, and pernicious weeds, Brassicaceae is a particularly well-studied family with broad comparative studies ranging from leaf morphology to life history (reviewed in [5]). Brassicaceae is also a model for comparative genomics with at least 21 published genomes [6] (<https://bmap.jgi.doe.gov>). By contrast, fewer species of Cleomaceae are economically important, with notable exceptions being the ornamental species, *Tarenaya hassleriana* [7] and the minor vegetable crop, *Gynandropsis gynandra* [8]. The floral ground plan of Brassicaceae is remarkably stable despite the size of the family, with little variation in arrangement of four-merous perianth and six stamens, whereas fruit morphology is extremely diverse. By contrast, flowers of Cleomaceae exhibit substantive diversity in symmetry, organ number, and elaboration of structures, whereas fruit morphology is relatively similar across most species. Thus, the contrast between these families represents an excellent model in which to investigate rates of lineage accumulation and the mechanisms underlying morphological diversification.

Research on Cleomaceae nicely complements the intense focus on Brassicaceae because Cleomaceae houses morphological and physiological diversity that is not present in Brassicaceae. Furthermore, the close phylogenetic distance between Cleomaceae and arabidopsis makes analysis of the underlying differences more tractable than for other plant groups, thereby facilitating the transfer of knowledge about development and genetic

Highlights

As broadening the comparative landscape becomes increasingly important, Cleomaceae emerges as a valuable plant model for groundbreaking inquiries that reflect its genomic, morphological, and physiological diversity, especially when compared to sister family the Brassicaceae.

Robust phylogenetic hypotheses are indispensable for providing an evolutionary comparative framework and structure needed for taxonomic revisions that impact on research ranging from genomics to physiology.

A genome triplication is a potential driver of floral evolution as well as a powerful system in which to explore the consequences of increased genome size.

Investigations in C₄ photosynthesis are instrumental in untangling the origin and maintenance of this ecologically relevant trait.

Remarkable floral diversity is an underutilized system to address the evolution of features associated with the important phenomenon of pollination.

¹Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

²RG Plant Cytogenomics, Central European Institute of Technology, 625 00 Brno, Czech Republic

³National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic

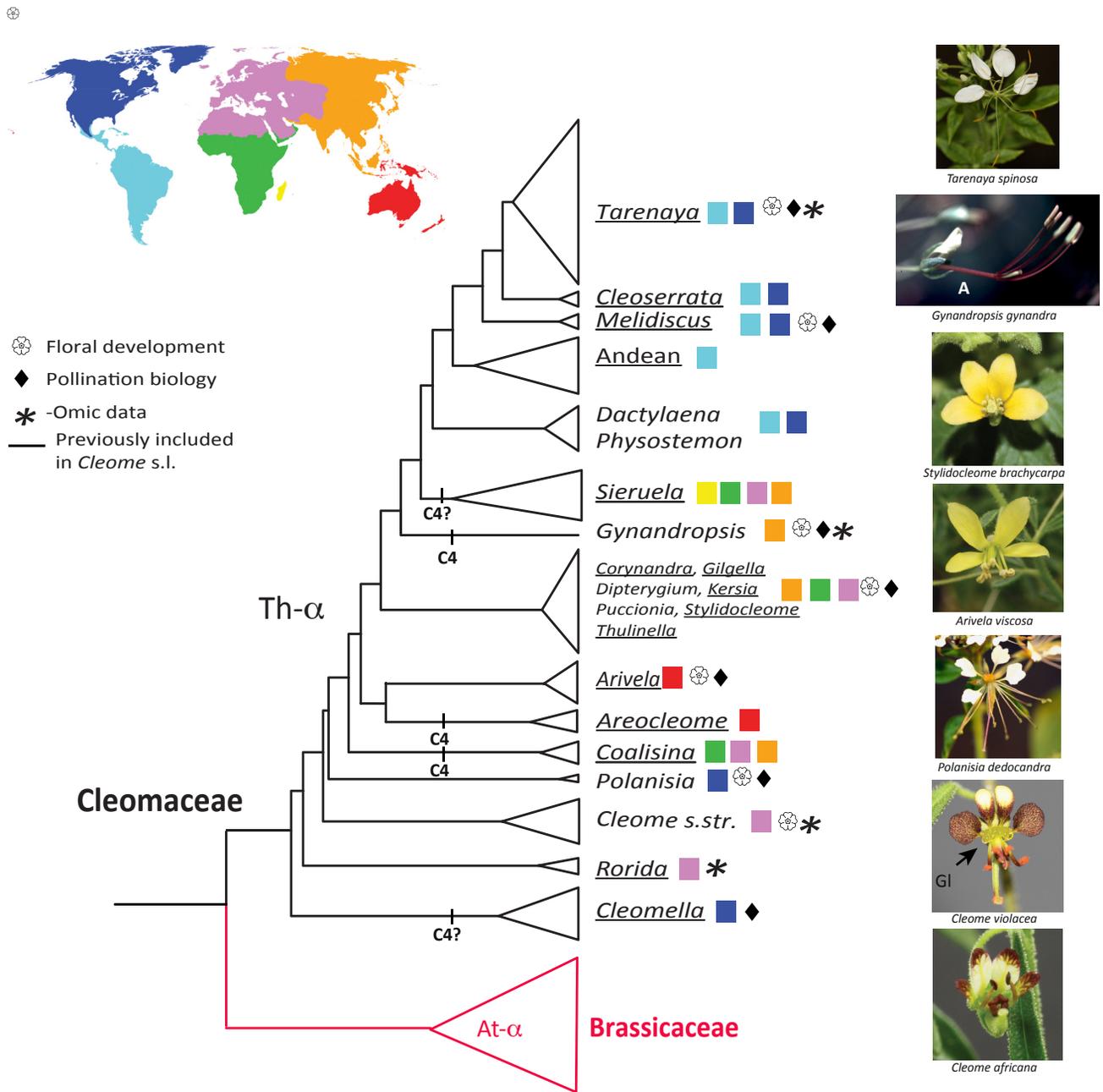
⁴Biosystematics Group, Wageningen University, 6708 PB Wageningen, The Netherlands

⁵School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA

*Correspondence: jocelyn.hall@ualberta.ca (J.C. Hall).

Key Figure

Relationships and Diversity of Cleomaceae



Trends in Plant Science

Figure 1. Fifteen principal clades have been identified that correlate more with geographic distribution than with the traditional boundaries of genera. Traditionally circumscribed *Cleome* species are distributed across multiple clades, necessitating a new classification of genera (genera that recently housed *Cleome* are underlined).

(Figure legend continued on the bottom of the next page.)

networks to relatives with novel and/or contrasting traits and expanding the comparative landscape around this premier model species. An essential requirement is a strong phylogenetic framework, which has recently been established for Cleomaceae. Further, shared and independent genome-duplication events between Cleomaceae and Brassicaceae establish the sister families as valuable models for studying gene and genome evolution after duplications, and how an increase in gene number may lead to the range of innovative characteristics that are present in these families [2,9,10]. Some of these traits have been extensively investigated (e.g., C_4 photosynthesis), whereas other areas are relatively unexplored as potential systems in which to address questions about plant diversity. We review here Cleomaceae as an emerging system to extend beyond the Brassicaceae in examining the origin of novelties and diversity, the consequences of gene duplications, and the genetic basis of adaptive traits.

A Strong Phylogenetic Framework for Cleomaceae Has Led to Recent Taxonomic Revisions

Because comparative studies rely on robust phylogenetic hypotheses and stable names, it is indispensable that recent systematic studies on Cleomaceae have made progress on both fronts [3,11–13]. Previously, 13 genera were recognized, with the vast majority of species being placed in the genus *Cleome*. Molecular phylogenetic studies have resulted in two major findings. First, *Cleome* as traditionally circumscribed is not monophyletic according to evaluations of nuclear, chloroplast, and mitochondrial markers [3,4,11,13–15]. These analyses established that *Cleome* is dispersed across many clades, which necessitates substantial nomenclatural revisions of the family (Figure 1, Key Figure; and Table 1). Second, 15 clades were identified based mostly on the most comprehensive study to date in which five molecular markers (*ndhF*, *matK*, *ycf1*, *rps3*, and ITS) were analyzed [13]. Instead of reflecting traditional generic boundaries, these clades are mostly correlated with geographical distributions (Figure 1). The family likely has an arid, western North American origin, with multiple dispersal events to Africa, although additional data and resolution are needed to confirm these historical biogeographic patterns [4].

In notable contrast to most molecular phylogenetic studies that uncover taxonomic problems but do not rectify them, the identification of well-supported clades within Cleomaceae has led to generic recircumscription. These revisions in turn ensure that researchers are making appropriate comparisons across clades versus erroneously assuming all traditionally circumscribed *Cleome* species are closely related (Figure 1). Taxonomic revision first focused exclusively on species traditionally placed in *Cleome* from the New World [16–20], with more recent revision of taxa and their distributions in the Old World [21–23]. The overall approach has been to combine molecular-based phylogenies with morphological patterns to guide nomenclatural changes. If one of the 15 principal clades correlates with morphological cohesiveness, then either resurrected or new names have been designated for one of the 15 principal clades (e.g., *Cleomella* [19], *Rorida* [22]). However, if one these clades includes morphologically heterogeneous taxa, then multiple genera have been described (e.g., clade 6 [13] was divided into seven genera: *Corynandra*, *Dipterygium*, *Gilgella*, *Kersia*, *Pucconia*, *Stylidocleome*, and *Thulinella* [21]; Figure 1). According to recent taxonomic

With the notable exception of the Andean clade, resurrected or new generic names are indicated at right, together with the geographical distributions of taxa within clades. Clades with published developmental (⊕), pollination (◆), and/or –omic (*) data are indicated (species and references are provided in Table 1). Known (C_4) and possible (C_4 -?) origins of C_4 photosynthesis are indicated on branches. Independent whole-genome duplication occurred within Brassicaceae (At- α) and Cleomaceae (Th- α), although the precise branch where Th- α took place is not known. Floral diversity is shown on far right, with androgynophore (A) and prominent floral nectary gland (G) as indicated.

Table 1. The 25 Recognized Genera of Cleomaceae^a

Genus	Number of species	Geographical distribution	Distinctive generic morphological traits	Focal species	Refs.
<i>Andinocleome</i>	~13	Predominantly Andean South America, with two species reaching Mexico, and one of these also occurs in Hispaniola and Venezuela	Small herbs to shrubs or small trees; leaves usually 5–13-foliolate; the nectiferous disc is usually conical or obconical and prominent; seeds usually lack an aril, uniquely winged in <i>C. stylosa</i>	<i>A. pilosa</i> ⊗-d	[59]
<i>Areocleome</i>	1+	Northern Australia	Purple flowers; leaves and stamen glabrous or with glandular hairs; stigma capitate; (6–)8–10 stamens; staminal filaments are dilated towards the apex	<i>A. oxalidea</i> C ₄	[42,55]
<i>Arivela</i>	12+	With the exception of <i>Arivela viscosa</i> , which has a worldwide distribution, the remaining species are distributed exclusively in Australia	Yellow flowers; stigma linear; stamens (1) 3–6, but occasionally as many as 52 in <i>A. viscosa</i> ; staminal filaments are not dilated towards the apex	<i>A. viscosa</i> ⊗-d, ◆	[59,66]
<i>Cleome</i> s.str.	27	Eastern Mediterranean, Spain, north and east Africa, Arabian Peninsula, and east to Afghanistan, Pakistan, Turkmenistan, and Kazakhstan	Annual or perennial, and then often woody at the base; leaves are unifoliolate or trifoliolate; flowers are zygomorphic, typically small to medium, and often with petals conspicuously veined; with stamens without staminodes	<i>C. violacea</i> ⊗-m, * ⊗-m <i>C. arabica</i> ⊗-m <i>C. ambylocarpa</i> ⊗-m	[59]
<i>Cleomella</i>	22 (25)	From central and western North America to south and central of Mexico	Small to medium sized, annual and predominantly glabrous; compound leaves with three entire leaflets; thick leaflets and often glaucous; filiform to lacinate stipules at the base of the petiole or pedicel; six stamens inserted on the androgynophore; oblong, cylindrical to rhomboidal capsules; seed claws are fused internally	<i>C. lutea</i> ◆ <i>C. serrulata</i> ◆ <i>C. sparsifolia</i> C ₄ ?	[55,64,65]
<i>Cleoserrata</i>	5	Mexico, Central and South America, West Indies	Glabrous annual herbs without spines or thorns; leaf margins are minutely serrulate; flowers are showy pink to purple or ivory or white		
<i>Coalisina</i>	6 (11)	Egypt, Yemen, Ethiopia, Namibia, South Africa, Angola, Sudan, Mozambique, India	Annual to short-lived perennial herbs, glabrous, moderately glandular pubescent, or covered in small prickles; 3–9(–13)-foliolate, linear or filiform leaflets; yellow flowers sometimes with purple markings; dimorphic stamens with most reduced to staminodes	<i>C. angustifolia</i> C ₄ <i>C. paradoxa</i> C ₃ –C ₄ /C ₂	[42,55]
<i>Corynandra</i>	5	From India to Malaysia	Leaves and stems with short, appressed hairs; purple flowers; stamens 6–250, staminal filaments dilated towards the apex; stigma linear		
<i>Dactylaena</i>	6	Argentina, Brazil, Haiti	Four linear petals; one fertile and four sterile stamens		
<i>Dipterygium</i>	1	Egypt to Pakistan	Low woody shrubs, glabrous to glandular scabrid; simple, elliptical oblong or ovate leaves; yellow flowers; six stamens; winged, one-seeded nut		
<i>Gilgella</i>	1	North and tropical Africa, Arabian Peninsula, east to Pakistan	Simple leaves; scapose inflorescence; lack of elongated and persistent style in fruit		
<i>Gynandropsis</i>	1	Asia	Erect annual herb with torus base which is subhemispherical; compound leaves are mostly penta-foliolate; white or pale pink flowers	<i>G. gynandra</i> ⊗-d, ◆, *, C ₄ , T	[36,42,44,46,52,59,66,68]
<i>Haptocarpum</i>	1	Brazil (Bahia)	Two functional stamens; fruits at dehiscence abscising the replum, leaving a forked base		
<i>Kersia</i>	8 (10)	Namibia, Angola, South Africa			

Table 1. (continued)

Genus	Number of species	Geographical distribution	Distinctive generic morphological traits	Focal species	Refs.
			Obovate to oblanceolate leaflets; yellow flowers; 12–35 dimorphic functional stamens		
<i>Melidiscus</i>	~2–4	Mexico to Brazil and Argentina	Typically shrubs 2–6 m tall; leaves 5–11-foliolate; petals usually green to yellow	<i>M. gigantea</i> ⊗-d, ◆	[59,67]
<i>Mitostylis</i>	2 (5)	Almost exclusively Caribbean, with one species also occurring in Belize	Perennials or annuals, usually decumbent to ascending; leaves unifoliolate; six stamens without apophyses		
<i>Physostemon</i>	10 (13)	Mexico to South America, with a few taxa in the Caribbean	Perennials or annuals, usually upright; leaves uni-, tri-, or pentafoliolate; stamens 6–10, most species have some stamens with apophyses		
<i>Podandrogynae</i>	~36	Southern Mexico to Bolivia; primarily Andean	Unisexual flowers (monoecious plants); petals often orange or orange-red; placentas contorted and seeds have a large aril		
<i>Polanisia</i>	6 (8)	North America, Mexico	Notched petals; large adaxial nectar glands; more than six stamens	<i>P. dodecandra</i> ⊗-m, ◆	[59,64]
<i>Puccionia</i>	1	Somalia	Cespitose herb, often woody at the base; glandular, simple ovate to round leaves; yellow flowers; six stamens; globose, rugose reticulate one-seeded nut		
<i>Rorida</i>	12	From north-eastern Africa through the Arabian Peninsula to Afghanistan, Tajikistan, and Pakistan	Simple leaves; four stamens; triangular, very short clawed, dimorphic petals with appendages at the base	<i>R. droserifolia</i> *	[69]
<i>Sieruela</i>	36 (40)	Kenya, Madagascar, Mozambique, Tanzania, South Africa, Southwest Africa, Ethiopia, Angola, Somalia, Congo, Sierra Leone, India	Annual to perennial erect herbs, occasionally woody; glabrous to glandular puberulent/glandular hirsute; 3–9-foliolate, with ovate, obovate, or oblong leaflets; purple flowers often with yellow markings on upper petals	<i>S. gallaensis</i> C _{4?} <i>S. allamanii</i> C _{4?}	[55]
<i>Stylidocleome</i>	1	Tropical and northeast Africa, east through the Arabian Peninsula, southern Iran, Afghanistan, Pakistan, and northwest India	Compound leaves; elongated and persistent style in fruit	<i>S. brachycarpa</i> ⊗-d	[59]
<i>Tarenaya</i>	30	Northern and eastern South America, north to southern Mexico, and the West Indies	Stipules are present as paired spines in most species; disks obsolete or disciform; most of the length of the mature seed cleft is covered by a fragile membrane	<i>T. hassleriana</i> ⊗-d, *, T <i>T. siliculifera</i> C _{4?}	[7,35,36,53,55,59,69]
<i>Thulinella</i>	1	Northeastern Africa, Arabian Peninsula, and east to southern Iran	Simple leaves; yellow petals with glandular appendages; 10–14 stamens; deflexed, contorted, persistent fruiting pedicels		

^aDetails include species number, geographical distribution, distinctive morphological traits, and species for which there are published -omic (denoted by *), pollination (denoted by ◆), and/or developmental data (denoted by ⊗), and whether the early developmental pathway is dissymmetric, -d, or monosymmetric, -m [59]. Species with full C₄ photosynthesis or some intermediate to C₄ photosynthesis are shown by C₄ and C_{4?}, respectively. Species that have established transformation protocols are indicated by T.

revisions, Cleomaceae currently comprises 25 genera (Table 1), with more genera to be described (e.g., Andean clade *sensu* [13]). Further, several genera/taxa have yet to be sampled in a molecular phylogenetic study [*Haptocarpum bahiense* (monotypic genus), *Mitostylis* spp., and most of the species diversity of *Physostemon*], and the relationships among lineages/genera are not yet resolved (e.g., Andean clade [8]). There is also a significant amount of new species discovery, which may impact on our understanding of lineage diversity [24].

Shared and Independent Genome Duplications within Cleomaceae and Brassicaceae Are Potential Drivers of Morphological Evolution

Genome duplication is a known driver of angiosperm diversification [25,26] and within the order Brassicales [4,27,28]. The Cleomaceae and Brassicaceae lineages have undergone both shared and independent WGD events. In addition to ancient duplications shared by seed plants, angiosperms, and eudicots [29], the Brassicales have experienced a β event shared by most families except Caricaceae, Moringaceae, Tropaeolaceae, and Akaniaceae [27,28]. Brassicales are united by the presence of secondary metabolites, called glucosinolates, which are antiherbivore compounds. In Brassicales, WGD events are correlated with increased rate of plant speciation and are followed by the innovation of novel glucosinolates [4,27]. In turn, these events are matched with gene duplication of counter-adaptations and codiversification of their main herbivores, pierid butterflies [27]. Combined, these patterns establish an arms race driven by WGD of plants and their herbivores. The resolution of relationships within Brassicales, which lead to appropriate placement of WGD events and the ability to assess increases in speciation rates, was accomplished by complementary phylogenomic approaches sampling one representative per family [27,28], as well as more intensive taxon sampling of fewer genes [4].

Additional WGD events occurred independently in Cleomaceae and Brassicaceae, named Th- α and At- α based on their documented occurrence in *Tarenaya hassleriana* and arabidopsis, respectively [7,10,30,31]. These independent events permit exploration of the consequences and evolutionary dynamics resulting from increases in genome size, and of whether an increase in gene number contributes to evolution of key traits that vary between the two families. Diploidization, through which a polyploidy genome transforms into a diploid one [31,32], is the most common fate of duplicated genes. However, some duplicated genes are maintained, often via either subfunctionalization or neofunctionalization, leading to increases in gene numbers and functions, and consequently to functional innovations [30,33,34]. The global retention rate of gene copy number is higher in Cleomaceae (50.0%) than in arabidopsis (13.6%), although this difference may simply reflect that *Tarenaya* experienced a hexaploid event more recently than the Brassicaceae duplication [2]. At- α occurred at 34 Mya, whereas the estimated date of Th- α is 20 Mya [30,31]. Brassicaceae houses much greater diversity in glucosinolates than Cleomaceae, and this correlates with retention of a greater number of genes involved in this pathway (49.1%), but Th- α did have a significant impact on glucosinolate diversity in Cleomaceae [2].

Given that Cleomaceae has a high global gene retention rate combined with greater floral variation, researchers have explored whether enrichment in particular gene families is associated with morphological diversity. This work has been based on the availability of the full

Box 1. C₄ Photosynthesis

The carbon-concentrating mechanism known as C₄ photosynthesis is a remarkable parallel innovation that is found in diverse flowering plants, including several species of Cleomaceae. Because the concentration of CO₂ increases around RuBisCo in C₄ photosynthesis (Figure 2), this altered pathway is an efficient mechanism to prevent photorespiration and to increase the photosynthetic activity of plants that are distributed in dry regions [40,46,48]. Thus, C₄ plants have a fitness advantage regarding carbon-fixation efficiency in under specific conditions (e.g., warm, sunny, and dry regions), and many important crop species are C₄ (e.g., corn [77]). Concentration of CO₂ is achieved via compartmentalization of carbon fixation (Figure 2). Instead of directly fixing CO₂ into a three-carbon compound, as in most angiosperms (called C₃ photosynthesis), C₄ species first convert CO₂ to four-carbon compound in mesophyll cells (Ms). These acids are then diffused to bundle-sheath cells (BSCs) to be fixed by RuBisCo. Two anatomical adaptations are observed in the leaves of C₄ plants: Kranz anatomy, in which one layer of M surrounds the BSC, and higher vein density [42,78]. Investigations often compare C₄ plants to C₃ plants to elucidate how this pathway may have evolved because extant C₃ plants are presumed to represent the ancestral state, and some studies have the long-term goal of modifying current C₃ crop species into C₄ [56,79].

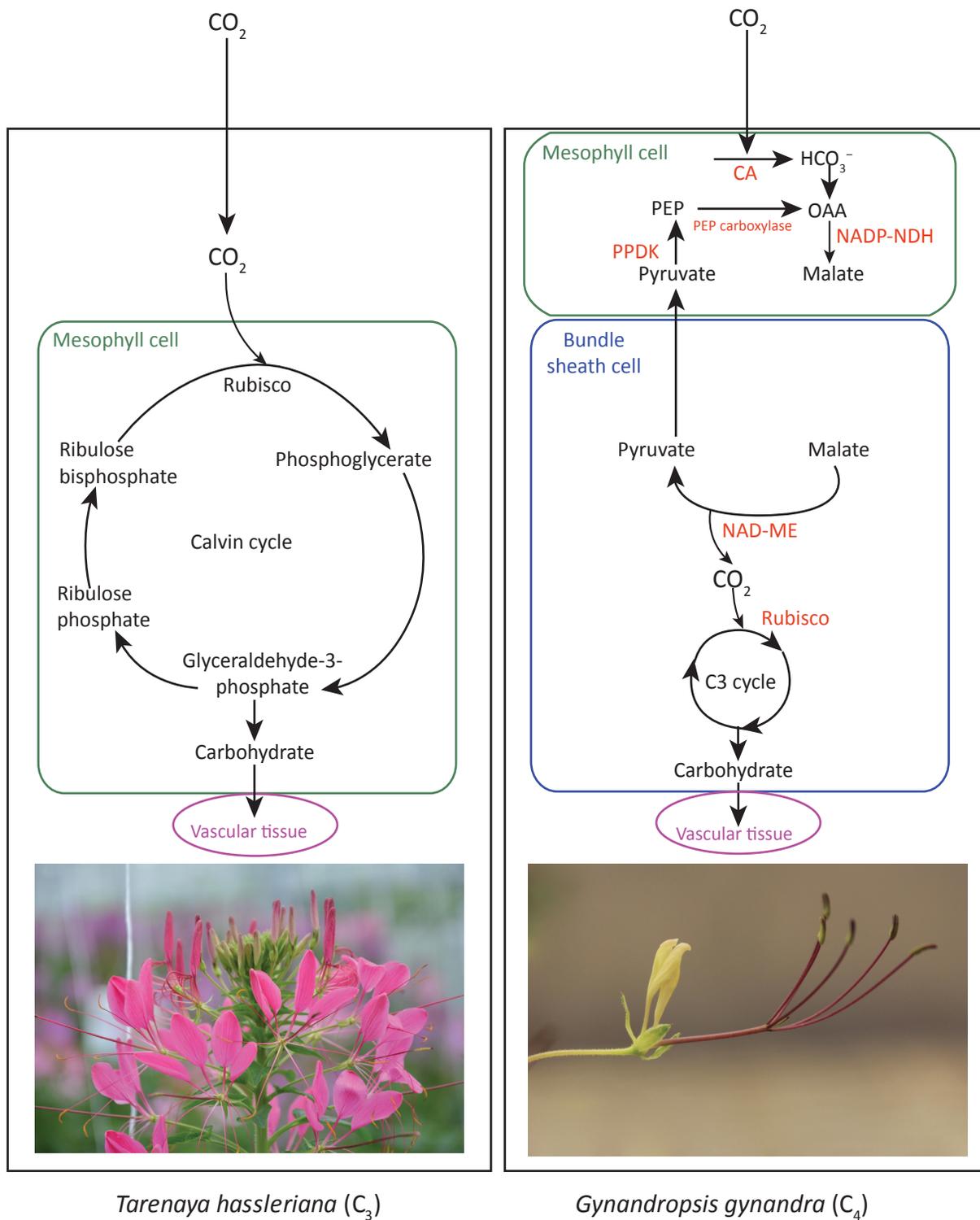
genome of *T. hassleriana* [7], as well as on transcriptomic libraries from this and other species [35,36]. For example, *TCP1* plays a crucial role in the transition from polysymmetric (multiple planes of symmetry) to monosymmetric (single plane of symmetry) flowers [37,38], and multiple copies are often associated with evolution of monosymmetry (reviewed in [38]). Disymmetry (two planes of symmetry)-dominated Brassicaceae retain one copy of *TCP1* [39], whereas there are at least two copies of *TCP1* in at least some Cleomaceae, which has predominantly monosymmetric flowers [7]. Similarly, Cleomaceae retains more copies of MADS-box genes that are associated with floral organ identity based on the ABC model. Specifically, it retains two copies of B-class genes *PISTILLATA1* (*PI*) and *APETALA3* (*AP3*), which are key regulators of stamen and petal identity in Arabidopsis, where only one copy of these genes has been retained in Brassicaceae [7]. The tandem duplication *AP3* in *Tarenaya* has also been hypothesized as a potential contributor to monosymmetry [7]. By contrast, the copy number of *SHATTERPROOF* (*SHP1/2*) genes that contribute to fruit morphology is higher in Brassicaceae than Cleomaceae. While there are four copies of *SHP1/2* in Brassicaceae, only one copy of this gene has been found in Cleomaceae, which coincidentally exhibit far less variation in fruit form than Brassicaceae [7].

Evolution of the C₄ Photosynthetic Pathway in Cleomaceae

C₄ photosynthesis has arisen >60 times independently in 19 different families including Cleomaceae (Box 1 and Figure 2) [10,40–43]. This remarkable convergence permits fundamental questions to be addressed on how this complex and ecologically important trait originated, and whether it evolved via the same route in independent plant lineages. Developmental shifts associated with shifts from C₃ to either C₃–C₄ intermediate/C₂ photosynthesis or full C₄ evolution include larger bundle-sheath cells (BSCs), increased leaf venation, more positive carbon isotope ratios, and fewer mesophyll cells (Ms) between veins. Accompanying such anatomical changes are clear shifts in expression patterns of photosynthesis-related genes to restrict key enzymes to vein, bundle-sheath, and/or mesophyll cells.

Within Cleomaceae, full C₄ photosynthesis has arisen at least three times [3], with other instances (at least three) of some deviation from what we expect for C₃ species (Figure 1). All three instances of C₄ in the family are of the NAD-ME subtype [42]. By far the most studied species is the African and Asian crop *Gynandropsis gynandra* [3,10,36,44–51]. This C₄ species is often compared to C₃ species *Tarenaya hassleriana* to elucidate the phenomenon of how C₄ photosynthesis evolved and is maintained [10,36,44]. Both species share the Th- α WGD event, and there is no significant difference in retained gene copy number between them [10], suggesting that differences in the timing and quantity of gene expression are responsible for the transition to C₄, and not the presence of additional gene copies. Valuable and detailed gene expression atlases were generated across multiple developmental stages and tissue types in both *G. gynandra* and *T. hassleriana* [36], which have been used in subsequent analyses (e.g., [2,10]). Because transformation techniques have been developed in both *G. gynandra* [52] and *T. spinosa* [53], hypotheses of how particular genes contribute to C₄ photosynthesis can be directly tested [49].

Investigations in Cleomaceae have been key in identifying how the anatomical differences, the restriction of enzymes to specific cells, and shifts in gene expression have evolved during the transition to C₄ photosynthesis. All enzymes that operate in C₄ photosynthesis are present in C₃ plants. Thus, researchers focus on two fundamental questions: (i) how do expression patterns of these genes and others differ between C₄ *G. gynandra* and C₃ *T. hassleriana* or C₃ Arabidopsis, and (ii) how is the restriction of different parts of the photosynthetic pathway between M and BSC achieved?



Tarenaya hassleriana (C₃)

Gynandropsis gynandra (C₄)

Trends in Plant Science

Figure 2. Comparison of Schematic Pathways of C₃ (Left) and C₄ (Right) Photosynthesis in *Tarenaya hassleriana* and *Gynandropsis gynandra*. Abbreviations: CA, carbonic anhydrase; NAD-ME, NAD malic enzyme; PPK, pyruvate phosphate dikinase; PEP carboxylase, phosphoenolpyruvate carboxylase.

Overall, gene expression profiles are similar between the two species, thus permitting pinpointing of differences in traits (e.g., photosynthetic pathways) that vary across them [36,45]. In other words, the modifications associated with C_4 leaves appear to have been built on a partly pre-existing regulatory architecture found in C_3 species. Perhaps unsurprisingly, different core C_4 genes are expressed at higher levels in *G. gynandra* than in *T. hassleriana* [36], notably within the guard cells that surround leaf stomata [44]. Temporal patterns explain some anatomical differences in the C_4 species *G. gynandra*. Delayed differentiation of C_4 leaves is correlated with a delayed onset of photosynthetic genes and genes that are involved in switching cell division to cell differentiation, which seems to promote the greater vein density observed in *G. gynandra* [36]. A master regulatory gene controlling differences between C_4 and C_3 gene expression has yet to be identified; however, promising candidates have been identified for future investigations [36,45]. Furthermore, growing understanding of the differences in expression profiles as a result of both *cis*-regulatory transcriptional and post-transcriptional processes has provided key new insights into the mechanisms underlying the profile differences [49,51].

Specifically, a question that has been addressed with Cleomaceae is how C_4 photosynthesis function is divided between BSCs and Ms, which requires that the expression of core genes of the C_4 cycle is restricted to specific cells (Figure 2). Focusing on carbonic anhydrases (CAs), which catalyze the first step of the C_4 pathway, site-directed mutagenesis analyses demonstrated that nine nucleotides (the MEM2 element) located in the 5' or 3' untranslated regions of the gene are sufficient to generate higher concentrations of CAs in M cells [49]. While the MEM2 element is crucial to M-specific expression, how this MEM2 element is recognized in C_4 leaves has yet to be determined, and determination of the unknown *trans*-acting factor will be necessary for understanding the evolution of this expression pattern. Importantly, however, researchers have demonstrated that alterations in *cis* and non-coding regions are a primary route through which cell-specific expression is achieved in *G. gynandra* [49,51].

A limiting factor of *G. gynandra* for studying fine-scale changes in the development of C_4 from C_3 is that it is an isolated and fully C_4 lineage, with no known close relatives that are C_3 [13,23]. Other Cleomaceae lineages need to be studied in more detail to rectify this. For instance, *Coalisina* [21] includes the C_4 *Coalisina angustifolia* and the C_3 - C_4 intermediate (or C_2 photosynthesis [54]) *C. paradoxa* [42]. However, based on carbon isotope ratios, at least three of the six species in this clade are C_3 [55], and there are six recognized subspecies and varieties of *C. angustifolia* [21], only one of which has been studied physiologically [42]. How much of this transition from C_3 to C_4 involves C_2 as an intermediate transitional step, as opposed to independent transitions in pathway optimization, is unclear without more detailed phylogenetic and physiological studies. There are also at least three other origins of modified photosynthetic pathways, at least as inferred by carbon isotope ratios: *Cleomella sparsifolia*, *Tarenaya siliculifera*, and *Sieruela allamanii* and *S. gallaensis* [55].

Recently, extensive variation within genotypes of *Gynandropsis gynandra* from Asia and Africa was documented [56]. Specifically, significant differences in vein density, cross-sectional area of BS strands, size of individual BS cells, stomatal density, photosynthetic performance, water-use efficiency, and gene expression patterns were found between genotypes [56]. Thus, future genetic mapping and/or association studies in Cleomaceae are possible to identify causal genes underlying standing natural variation in the components of C_4 photosynthesis.

In sum, researchers are identifying sets of genes whose expression patterns vary in C_4 plants, and are determining the mechanisms by which these differences are achieved between species and across different cell types. Further, findings in Cleomaceae can be compared to other C_4 species to understand commonalities and differences underlying this complex trait across independent origins. Different systems of C_4 photosynthesis have been described and compared, highlighting unique attributes across diverse taxa (recently reviewed in [48,50,57,58]), with *G. gynandra* emerging as extremely well-studied C_4 species, but with a need for more studies in the other C_4 lineages of the family.

Morphological Variation across Cleomaceae Represents Untapped Systems in Which To Address Fundamental Evolutionary Questions

Cleomaceae house exquisite morphological variation, especially in floral features, that aid in investigations of the evolution of traits associated with pollination attraction. The majority of flowers in Cleomaceae are monosymmetric, which is imposed via upward curvature of petals and reproductive organs [59] (Figure 1). Monosymmetry is generally touted as an adaptive innovation that leads to rampant speciation, as evidenced by many of the largest plant families having monosymmetric flowers (e.g., Orchidaceae, Fabaceae) [60–62], but it is notable that the number of Cleomaceae species is one tenth of that of disymmetric Brassicaceae. The increase in speciation and monosymmetry is often associated with specialized pollination syndromes [63]. Although there have been few pollination biology studies of Cleomaceae, they also show that these species are the exception to the rule. Five species are generalists based on empirical studies [64–66], whereas others are described as being pollinated by either bats or hawkmoths [67,68].

Pollination studies are complemented by the description of two distinct developmental pathways that produce similar forms of mature monosymmetric flowers [59]. In some species the adaxial sepal completely envelops the developing bud (termed early monosymmetry). By contrast, immature sepals of other species are of equal size and shape during development, and this was termed the early disymmetric pathway. The early disymmetric pathway appears to be derived within Cleomaceae (Table 1), and is accompanied by similarities in the mature adaxial and abaxial petals. The genetic basis and whether developmental pathways constrain mature floral morphology remain understudied, as does whether this pathway is associated with the WGD event (*G. gynandra* and *T. hassleriana* share Th- α and exhibit the derived disymmetric pathway). Thus, developmental and pollination work has been completed on some of the same species (Table 1), leading to a range of questions including whether there are undescribed patterns of floral development, what are the genetic bases of monosymmetry, and how monosymmetry may impact on pollination attraction. Detailed transcriptomics from a range species provide rich sources of data to initially explore these questions [35,36,69], as demonstrated by the use of one dataset to examine whether glucosinolate diversity differences between Brassicaceae and Cleomaceae may play a role in bat pollination attraction [2].

While many flowers in Cleomaceae have similar symmetry at maturity, variation is observed in other floral traits, permitting investigations of additional features associated with pollinator attraction and rewards. Petal and sepal size vary within a flower and across species of different clades. For instance, *Tarenaya hassleriana* and *Arivela viscosa* have equally sized petals, whereas *Cleome violacea* and *C. africana* have unequally sized petals (Figure 1). Differences in flower petal size in *Iberis amara* (Brassicaceae) are driven by increased *TCP1* expression [37,39], which raises the question of whether the same or alternative mechanisms drive petal-size differences in Cleomaceae. Perhaps more directly related to pollination attraction, petal color ranges from pink to purple, or yellow to white or dull greenish, or orange to red [1]

(Table 1). The pink-purple color in petals of *Tarenaya hassleriana* and *Cleomella serrulata* is produced by anthocyanin pigments [70]. Differences in anthocyanins, and in subsequent petal color, are an adaptive trait for pollinator attraction and have been identified as an important but understudied component of evolutionary/developmental studies that are in need of additional research across taxa [71,72]. Further, known regulators of anthocyanin biosynthesis genes have shown to be upregulated in pink-flowered *T. hassleriana* [35], and these data could be compared to different-colored flowers in other species. Further, petal spots are observed in some Cleomaceae species. For instance, in *Cleome violacea* yellow spots are present on adaxial petals, but abaxial petals are uniformly maroon (Figure 1). These petal spots are also seen in *Cleome africana* and *Styldocleome brachycarpa*; however, the position of these spots is different compared to *C. violacea* [59]. This observed variation permits inquiries within and between clades to test whether these adaptive traits have the same or different genetic bases as other plants.

While petal size and petal color have been studied in other systems, other traits putatively connected to pollinator response in Cleomaceae have not been as intensely investigated [7,59,64]. For example, the substantive variation in size, prominence, and output of nectary glands across the family likely impacts on variation in pollinators [73], and has not been explored in a phylogenetic context. Further, the genetic basis of nectary glands has been well studied in *Arabidopsis* and relatives [74–76], although this species is predominantly selfing. This information facilitates exploring how differences in nectary glands are generated as well as whether the genetic mechanisms of gland formation are convergent or divergent with Brassicaceae. Similarly, almost all Cleomaceae species exhibit stalks or stems to the ovary, called gynophores, whereas some species have staminal stalks, called androphores. Not only does the length of gynophores vary, these stalks may be fused with androphores to generate combined structures, called androgynophores, a structure that is notably found in the well-investigated species *G. gynandra* (Figure 1). These stalks alter the height and position of stigma and anthers, and this likely directly impacts on efficacy of putative insect visitors. Despite this likely connection to pollination biology, and the presence of androgynophores among taxa within Cleomaceae and beyond (e.g., *Passiflora*), the genes responsible for androgynophore development have not yet been characterized. Knowledge of these genes could potentially be used to modify the crop plant *G. gynandra* to attract more pollinators, increasing the seed stock for rural communities that rely on this crop for food security.

Morphological variation within Cleomaceae extends beyond the flowers, including variation in features such as growth form (woody versus herbaceous). For example, there are likely multiple origins of woody perennials in the family (e.g., *Andiocleome*, *Cleomella*, *Melidiscus*, *Tarenaya*). Given the propensity for the family to grow in arid environments, combined with multiple origins of C_4 photosynthesis, Cleomaceae is a promising group to correlate the presence of C_4 photosynthesis with the evolution of growth forms to provide greater ecological context to the evolution of this alternative pathway. To date, no study has compared the growth rate (e.g., the growth rate of roots compared to that of the stem or the rate of producing leaves) of C_3 and C_4 species in Cleomaceae. Moreover, few studies have focused on the genetic mechanisms that may alter the growth form in Cleomaceae, how these forms are integrated with C_4 , and the general evolution of these traits. Despite poor knowledge about the effect of C_4 photosynthesis on leaf morphology and growth form evolution of C_4 Cleomaceae species, leaf anatomical adaptations of C_4 species have been well documented [42]. *Cleome angustifolia* shows other leaf morphological adaptations, such as decreased leaf area, the presence of water-storage tissues, and a decrease in reticulate leaf venation, that are characteristics of plants growing in extremely dry areas.

Concluding Remarks and Future Perspectives

Recent advances in sequencing, genome engineering, and tools permit assessment of gene function in a range of species, thereby enabling greater species diversity to be used in comparative evolutionary studies. It is clear that our understanding of the evolution and diversity of plants can be improved when a broader range of species are studied, especially those closely related to model species where transfer of knowledge is facilitated. Cleomaceae represents such a system, and houses novelties and presents a notable contrast to Brassicaceae with regard to floral diversity, gene retention, and species diversification. Research in Cleomaceae suggests that some morphological diversity, including variation in floral form, may be driven by gene copy-number increase, whereas other traits, notably C_4 photosynthesis, are the result of complex differences in the expression patterns of many genes. These trends, combined with robust phylogenetic hypotheses and subsequent taxonomic revisions, provide a strong context for further evaluation of diversification patterns and their underlying processes in Cleomaceae (see Outstanding Questions).

Acknowledgments

S.B. is supported by the Natural Science and Engineering Research Council of Canada (NSERC). We thank members of the laboratories of J.C.H., M.E.S., and E.H.R. for many fascinating discussions on Cleomaceae. Brandi Zenchyzen kindly provided the photo of *Gynandropsis gynandra*. We thank Frank Becker for the photograph of *Tarenaya hassleriana* in Figure 2.

References

1. Iltis, H.H. *et al.* (2011) Studies in the Cleomaceae I: on the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae. *Ann. Missouri Bot. Garden* 98, 28–36
2. van den Bergh, E. *et al.* (2016) Flower power and the mustard bomb: comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *Am. J. Bot.* 103, 1212–1222
3. Feodorova, T.A. *et al.* (2010) Biogeographic patterns of diversification and the origins of C_4 in *Cleome* (Cleomaceae). *Syst. Bot.* 35, 811–826
4. Cardinal-McTeague, W.M. *et al.* (2016) Biogeography and diversification of Brassicales: a 103 million year tale. *Mol. Phylogenet. Evol.* 99, 204–224
5. Koenig, D. and Weigel, D. (2015) Beyond the thale: comparative genomics and genetics of *Arabidopsis* relatives. *Nat. Rev. Genet.* 16, 285–298
6. Nikolov, L.A. and Tsiantis, M. (2017) Using mustard genomes to explore the genetic basis of evolutionary change. *Curr. Opin. Plant Biol.* 36, 119–128
7. Cheng, S. *et al.* (2013) The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* 25, 2813–2830
8. Sogbohossou, E.O.D. *et al.* (2018) A roadmap for breeding orphan leafy vegetable species: a case study of *Gynandropsis gynandra* (Cleomaceae). *Hortic. Res.* 5, Published online January 10, 2018. <http://dx.doi.org/10.1038/s41438-017-0001-2>
9. Marshall, D.M. *et al.* (2007) *Cleome*, a genus closely related to *Arabidopsis*, contains species spanning a developmental progression from C_3 to C_4 photosynthesis. *Plant J.* 51, 886–896
10. van den Bergh, E. *et al.* (2014) Gene and genome duplications and the origin of C_4 photosynthesis: birth of a trait in the Cleomaceae. *Curr. Plant Biol.* 1, 2–9
11. Hall, J.C. (2008) Systematics of Capparaceae and Cleomaceae: an evaluation of the generic delimitations of *Capparis* and *Cleome* using plastid DNA sequence data. *Bot. Bot.* 86, 682–696
12. Hall, J.C. *et al.* (2004) Molecular phylogenetics of core Brassicales, placement of orphan genera *Emblingia*, *Forchhammeria*, *Triania*, and character evolution. *Syst. Bot.* 29, 654–669
13. Patchell, M.J. *et al.* (2014) Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* 63, 315–328
14. Hall, J.C. *et al.* (2002) Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. *Am. J. Bot.* 89, 1826–1842
15. Tamboli, A.S. *et al.* (2016) Phylogenetic analysis, genetic diversity and relationships between the recently segregated species of *Corynandra* and *Cleoserrata* from the genus *Cleome* using DNA barcoding and molecular markers. *C. R. Biol.* 339, 123–132
16. Cochrane, T.S. and Iltis, H.H. (2014) Studies in the Cleomaceae VII: five new combinations in *Corynandra*, an earlier name for *Arivela*. *Novon* 23, 21–26
17. 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. *Novon* 17, 447–451
18. Iltis, H.H. and Cochrane, T.S. (2014) Studies in the Cleomaceae VI: a new genus and sixteen new combinations for the Flora Mesoamericana. *Novon* 23, 51–58
19. Roalson, E.H. *et al.* (2015) A revision of generic boundaries and nomenclature in the North American cleomoid clade (Cleomaceae). *Phytotaxa* 205, 129–144
20. Neto, R.L.S. *et al.* (2017) *Cleoserrata* (Cleomaceae): taxonomic considerations and a new species. *Phytotaxa* 324, 179–186
21. Roalson, E.H. and Hall, J.C. (2017) New generic concepts for African Cleomaceae. *Syst. Bot.* 42, 925–942
22. Thulin, M. and Roalson, E.H. (2017) Resurrection of the genus *Florida* (Cleomaceae), a distinctive Old World segregate of *Cleome*. *Syst. Bot.* 42, 569–577
23. Barrett, R.L. *et al.* (2017) Resolving generic boundaries in Indian-Australasian Cleomaceae: circumscription of *Areocleome*, *Arivela*, and *Corynandra* as distinct genera. *Syst. Bot.* 42, 15
24. Soares Neto, R.L. *et al.* (2018) Two new species of *Tarenaya* (Cleomaceae) from Brazil. *Phytotaxa* 334, 028–034
25. De Bodt, S. *et al.* (2005) Genome duplication and the origin of angiosperms. *Trends Ecol. Evol.* 20, 591–597
26. Soltis, D.E. *et al.* (2009) Polyploidy and angiosperm diversification. *Am. J. Bot.* 96, 336–348

Outstanding Questions

What is the biogeographic history of Cleomaceae, and how might these patterns inform on morphological evolution, the origin of C_4 photosynthesis, and species diversification?

Where phylogenetically did the genome triplication occur within Cleomaceae, what is the impact of this increase in genetic inventory, and are the patterns and consequences documented in this family similar to or contrasted with parallel events in Brassicaceae?

Although not implicated in C_4 evolution, is retaining a high global gene copy number after whole-genome triplication a major contributor towards other morphological innovations (e.g., androgynophore, habit) and floral variation (e.g., alteration in the position of petal spots, variation in floral glands)?

Does the origin and maintenance of C_4 photosynthesis in *Gynandropsis* converge or diverge with other C_4 species in Cleomaceae and other angiosperm families, and how might this knowledge be applied to crop improvement and understanding the evolution of complex traits?

What is the genetic basis of floral traits associated with pollinator attraction and rewards in understudied tissues such as glands and androgynophores?

27. Edger, P.P. *et al.* (2015) The butterfly plant arms-race escalated by gene and genome duplications. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8362–8366
28. Edger, P.P. *et al.* (2018) Brassicales phylogeny inferred from 72 plastid genes: a reanalysis of the phylogenetic localization of two paleopolyploid events and origin of novel chemical defenses. *Am. J. Bot.* 105, 463–469
29. Jiao, Y. *et al.* (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100
30. Barker, M.S. *et al.* (2009) Paleopolyploidy in the Brassicales: analyses of the *Cleome* transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. *Genome Biol. Evol.* 1, 391–399
31. Schranz, M.E. and Mitchell-Olds, T. (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* 18, 1152–1165
32. Mandakova, T. *et al.* (2016) How diploidization turned a tetraploid into a pseudotriploid. *Am. J. Bot.* 103, 1187–1196
33. Conant, G.C. and Wolfe, K.H. (2008) Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* 9, 938–950
34. Rastogi, S. and Liberles, D.A. (2005) Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol. Biol.* 5, 28
35. Bhide, A.S. *et al.* (2014) Analysis of the floral transcriptome of *Tarenaya hassleriana* (Cleomaceae), a member of the sister group to the Brassicaceae: towards understanding the base of morphological diversity in Brassicales. *BMC Genomics* 15, 140
36. K ulahoglu, C. *et al.* (2014) Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* 26, 3243–3260
37. Busch, A. *et al.* (2014) Differential transcriptome analysis reveals insight into monosymmetric corolla development of the crucifer *Iberis amara*. *BMC Plant Biol.* 14, 285
38. Rosin, F.M. and Kramer, E.M. (2009) Old dogs, new tricks: regulatory evolution in conserved genetic modules leads to novel morphologies in plants. *Dev. Biol.* 332, 25–35
39. Busch, A. *et al.* (2012) Corolla monosymmetry: evolution of a morphological novelty in the Brassicaceae family. *Mol. Biol. Evol.* 29, 1241–1254
40. Brown, N.J. *et al.* (2011) Independent and parallel recruitment of preexisting mechanisms underlying C-4 photosynthesis. *Science* 331, 1436–1439
41. Koteyeva, N.K. *et al.* (2014) Differentiation of C-4 photosynthesis along a leaf developmental gradient in two *Cleome* species having different forms of Kranz anatomy. *J. Exp. Bot.* 65, 3525–3541
42. Koteyeva, N.K. *et al.* (2011) Diversity in forms of C-4 in the genus *Cleome* (Cleomaceae). *Ann. Bot.* 107, 269–283
43. Sage, R.F. *et al.* (2011) The C4 plant lineages of planet Earth. *J. Exp. Bot.* 62, 3155–3169
44. Aubry, S. *et al.* (2016) A specific transcriptome signature for guard cells from the C-4 plant *Gynandropsis gynandra*. *Plant Physiol.* 170, 1345–1357
45. Brautigam, A. *et al.* (2011) An mRNA blueprint for C-4 photosynthesis derived from comparative transcriptomics of closely related C-3 and C-4 species. *Plant Physiol.* 155, 142–156
46. Brautigam, A. *et al.* (2011) Critical assessment of assembly strategies for non-model species mRNA-Seq data and application of next-generation sequencing to the comparison of C-3 and C-4 species. *J. Exp. Bot.* 62, 3093–3102
47. Covshoff, S. *et al.* (2014) Getting the most out of natural variation in C-4 photosynthesis. *Photosynth. Res.* 119, 157–167
48. Williams, B.P. *et al.* (2012) Molecular evolution of genes recruited into C-4 photosynthesis. *Trends Plant Sci.* 17, 213–220
49. Williams, B.P. *et al.* (2016) An untranslated cis-element regulates the accumulation of multiple C-4 enzymes in *Gynandropsis gynandra* mesophyll cells. *Plant Cell* 28, 454–465
50. Huang, C.F. *et al.* (2016) Insights into the regulation of C-4 leaf development from comparative transcriptomic analysis. *Curr. Opin. Plant Biol.* 30, 1–10
51. Fankhauser, N. and Aubry, S. (2017) Post-transcriptional regulation of photosynthetic genes is a key driver of C-4 leaf ontogeny. *J. Exp. Bot.* 68, 137–146
52. Newell, C.A. *et al.* (2010) *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a C-4 dicotyledon that is closely related to *Arabidopsis thaliana*. *J. Exp. Bot.* 61, 1311–1319
53. Tsai, Y.T. *et al.* (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
54. Sage, R.F. *et al.* (2012) Photorespiration and the evolution of C-4 photosynthesis. *Annu. Rev. Plant Biol.* 63, 19–47
55. Voznesenskaya, E.V. *et al.* (2007) Physiological, anatomical and biochemical characterisation of photosynthetic types in genus *Cleome* (Cleomaceae). *Funct. Plant Biol.* 34, 247–267
56. Reeves, G. *et al.* (2018) Natural variation within a species for traits underpinning C4 photosynthesis. *Plant Physiol.* 177, 504–512
57. Burgess, S.J. and Hibberd, J.M. (2015) Insights into C-4 metabolism from comparative deep sequencing. *Curr. Opin. Plant Biol.* 25, 138–144
58. Schluter, U. *et al.* (2016) Understanding metabolite transport and metabolism in C-4 plants through RNA-seq. *Curr. Opin. Plant Biol.* 31, 83–90
59. Patchell, M.J. *et al.* (2011) Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *Int. J. Plant Sci.* 172, 352–365
60. Jabbour, F. *et al.* (2009) Evolution of floral symmetry: a state of the art. *C. R. Biol.* 332, 219–231
61. Hilleman, L.C. (2014) Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130348
62. Vamosi, J.C. and Vamosi, S.M. (2010) Key innovations within a geographical context in flowering plants: towards resolving Darwin's abominable mystery. *Ecol. Lett.* 13, 1270–1279
63. Armbruster, W.S. (2017) Plant pollinator interactions from flower to landscape: the specialization continuum in pollination systems: diversity of concepts and implications for ecology, evolution and conservation. *Funct. Ecol.* 31, 88–100
64. Higuera-Diaz, M. *et al.* (2015) Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* 93, 745–757
65. Cane, J.H. (2008) Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biol.* 23, 152–158
66. Raju, A. and Rani, D. (2016) Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytol. Balcan.* 22, 15–28
67. Fleming, T.H. *et al.* (2009) The evolution of bat pollination: a phylogenetic perspective. *Ann. Bot.* 104, 1017–1043
68. Martins, D.J. and Johnson, S.D. (2013) Interactions between hawkmoths and flowering plants in East Africa: polyphagy and evolutionary specialization in an ecological context. *Biol. J. Linn. Soc.* 110, 199–213
69. Mohammadin, S. *et al.* (2015) Positionally-conserved but sequence-diverged: identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biol.* 15, 12
70. Nozzolillo, C. *et al.* (2010) Novel aspects of the flowers and floral pigmentation of two *Cleome* species (Cleomaceae), *C. hassleriana* and *C. serrulata*. *Biochem. Syst. Ecol.* 38, 361–369
71. Sobel, J.M. and Streisfeld, M.A. (2013) Flower color as a model system for studies of plant evo-devo. *Front. Plant Sci.* 4, 321
72. Roberts, W.R. and Roalson, E.H. (2017) Comparative transcriptome analyses of flower development in four species of *Achimenes* (Gesneriaceae). *BMC Genomics* 18, 26

73. Bender, R. *et al.* (2012) Functional genomics of nectar production in the Brassicaceae. *Flora* 207, 491–496
74. Kram, B.W. *et al.* (2009) Uncovering the *Arabidopsis thaliana* nectary transcriptome: investigation of differential gene expression in floral nectariferous tissues. *BMC Plant Biol.* 9, 16
75. Thomas, J.B. *et al.* (2017) The pennycress (*Thlaspi arvense* L.) nectary: structural and transcriptomic characterization. *BMC Plant Biol.* 17, 9
76. Bender, R. *et al.* (2012) Functional genomics of nectar production in the Brassicaceae. *Flora* 207, 491–496
77. Moreno-Villena, J.J. *et al.* (2018) Highly expressed genes are preferentially co-opted for C-4 photosynthesis. *Mol. Biol. Evol.* 35, 94–106
78. Heckmann, D. (2016) C-4 photosynthesis evolution: the conditional Mt. Fuji. *Curr. Opin. Plant Biol.* 31, 149–154
79. Gowik, U. and Westhoff, P. (2011) The path from C₃ to C₄ photosynthesis. *Plant Physiol.* 155, 56–63