Transcriptional dissection of petty spurge (*Euphorbia peplus* L.) reproductive structures supports the *Euphorbia* cyathium as an inflorescence rather than a complete flower

Authors: Arielle Rose Johnson[1], Ashley Bao[1], Margaret Hannah Frank[1]
Affiliation: [1]Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, USA

Abstract: The unique reproductive structure of *Euphorbia* species, the cyathium, has historically been considered an inflorescence despite its visual resemblance to a single complete flower. However, multiple other models have recently been proposed, including the idea that the cyathium is a flower or has some degree of floral identity. Enabled by the petty spurge (*Euphorbia peplus* L.) genome and guided by the ABCDE model of floral development, we dissected petty spurge cyathia and carried out a transcriptomic analysis of the different organs. We also constructed gene phylogenies and performed dN/dS analysis on select floral genes. The E class gene SEP3 was not upregulated in the involucre as compared with the cyathophyll, indicating that the cyathium is not a complete flower. The filiform structures show upregulation of B and E class genes and transcriptomic signatures of heterochromatin formation, consistent with the hypothesis that they are reduced flowers. *E. peplus* homologs for the inflorescence/floral meristem genes UFO and LFY and the B class genes AP3 and PI have highly diverged sequences relative to other Euphorbiaceae species, suggesting that these reproductive patterning genes may have evolved divergent functions. We propose a new model in which the evolution of the cyathium involves two innovations: 1. altering LFY function to influence the timing of floral meristem development to create a compact flower-like inflorescence, and 2. altering the B class genes so that no perianth is formed. Our new hypothesis includes testable predictions about protein interactions and gene sequence evolution.

Key words: *Euphorbia peplus*, *Euphorbia*, cyathium, spurge, inflorescence

There are approximately 2000 species of *Euphorbia*, including horticulturally important plants such as poinsettia (*Euphorbia pulcherrima* L). The reproductive structure of *Euphorbia* species is called a cyathium. A cyathium looks similar to a complete flower with pistils, stamens, and a perianth; however, it has historically been interpreted as a pseudanthium (false flower) that is a condensed inflorescence consisting of a pistillate flower surrounded by staminate flowers and an involucral bract (Figure 1).

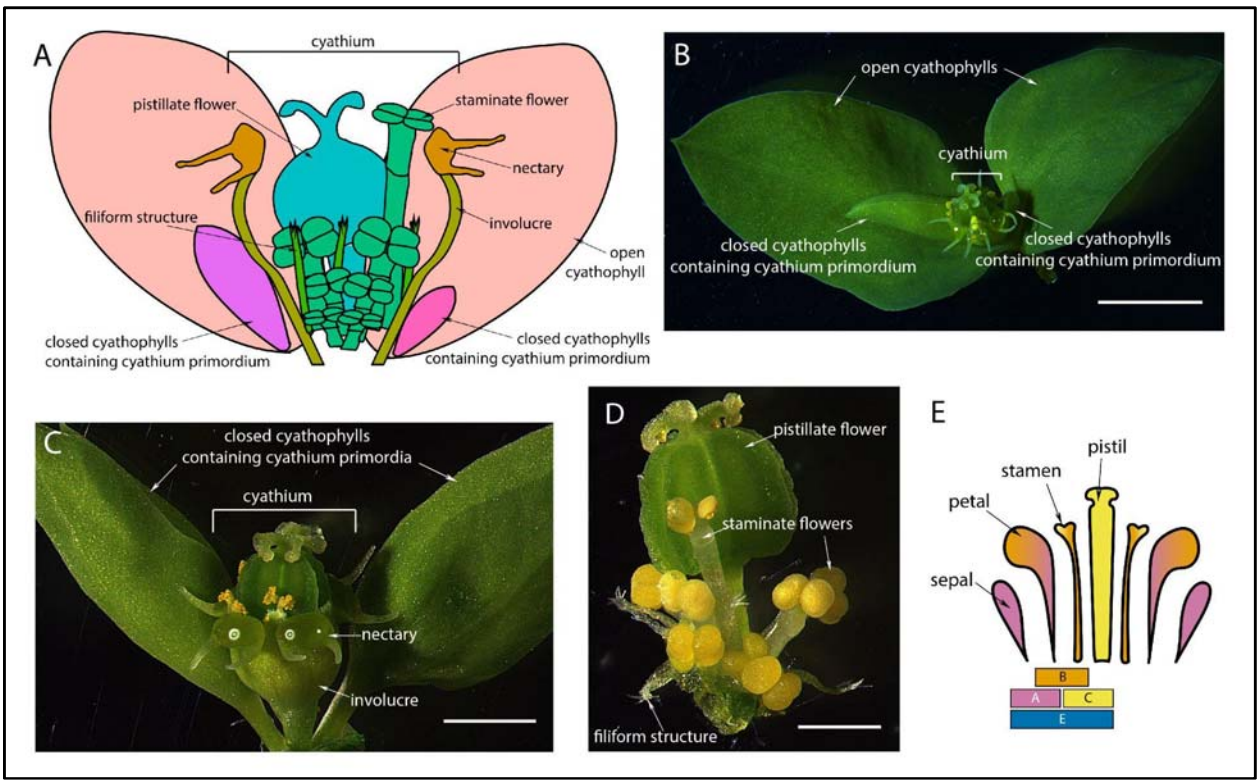


Figure 1

Euphorbia peplus L., commonly known as petty spurge, is an emerging model species for *Euphorbia* with a publicly available genome assembly (Czechowski et al., 2022; Johnson et al., 2023). In *E. peplus*, each immature cyathium is enclosed within two bracts, called cyathophylls. The cyathophyll contains the immature cyathium as well as multiple smaller cyathium primordia (enclosed in their own cyathophylls). The number of primordia varies by *E. peplus* ecotype. The

cyathophylls open at approximately the stage of stigma lobe morphology which corresponds to high stigma receptivity (Li et al., 2014), bringing the cyathium into the open for pollination and later seed dispersal. *E. peplus* is primarily self-pollinated.

The *E. peplus* cyathium itself, as in other *Euphorbia* species, is composed of a single pistillate flower in the center, surrounded by five mature staminate flowers. Each of the five mature staminate flowers is subtended by immature or nonfunctional staminate flowers in a zigzag pattern. Interspersed with the staminate flowers in the *E. peplus* cyathium are fringed filiform (threadlike) structures. These filiform structures have been referred to by various terms including “filamentous structures” (Prenner and Rudall, 2007) and “filiform bracteoles” (Bruyns, 2022). For the sake of clarity we will refer to these structures as “filiform structures" in this paper, because “filamentous” implies a developmental homology between these structures and filaments, and because Prenner and Rudall 2007 also refer to cyathophylls as “bracteoles” (Prenner and Rudall, 2007). The pistillate flower, staminate flowers, and filiform structures are surrounded by a fused cuplike structure of involucral bracts, also known as the involucre. It is this cup-shaped structure that gives cyathia their name, as “cyathium” comes from the Greek word for cup, “kyathos”. Crescent-shaped glands or nectaries are present near the top of the involucre. In *E. peplus*, there are 4 or 5 nectaries present; this number varies by species.

Historically, the cyathium has been interpreted as an inflorescence for several reasons that are well reviewed in Prenner and Rudall 2007 (Prenner and Rudall, 2007): the cyathium has a line of constriction on the staminate flower that is interpreted as the flower-stamen boundary; other genera in the Euphorbiaceae (e.g. *Anthostema*, Figure 2) have staminate flowers clustered around a pistillate flower; and flowers are usually unisexual in the Euphorbiaceae (Kaplan and Specht, 2022). Euphorbiaceae genera such as *Dalechampia* and *Pera* have reproductive structures that have long been established to be pseudanthia (Gagliardi et al., 2018), and

morphological work on which Euphorbiaceae have pseudanthia is still ongoing, e.g. recent work interpreting the “staminate flower” of *Ricinus communis* as a multi-flower unit(Claßen-Bockhoff and Frankenhäuser, 2020). Prenner and Rudall 2007 speculate that if the cyathium is an inflorescence, the filiform structures could be interpreted either as prophylls or as “highly reduced male flowers that represent remnants of the original dichasial male inflorescence. This would explain the variable number of [filiform] structures and the fact that they are not associated with particular male flowers”(Prenner and Rudall, 2007). The involucre would then be interpreted as a non-floral organ.

The Australian genus *Calycopeplus* and the New Caledonia tree *Neoguillauminia cleopatra* have been proposed as “missing links” towards the evolution of the *Euphorbia* cyathium from separate pistillate and staminate flowers; these close relatives to *Euphorbia* both have an inflorescence where a pistillate flower with a small perianth is surrounded by staminate flowers with small perianths (Figure 2) (Western Australian Herbarium et al., n.d.; Prenner and Rudall, 2007; Prenner et al., 2008; Dressler et al., 2014).

Calycopeplus paucifolius © overlander (Gerald Krygsman), iNaturalist

Neoguillauminia cleopatra © Benoit Henry, Pl@ntNet

Anthostema senegalense © Philippe Birnbaum, African Plants-- A Photo Guide

Figure 2

However, Prenner and Rudall posit that the case for *Euphorbia* as a pseudanthium is not clear-cut: other Euphorbiaceae have abscission zones that are not at the base of an organ; transitioning between an inflorescence arrangement like *Anthostema*'s and the inflorescence arrangement of the *Euphorbia* cyathium would require a large change in the orientation of the flowers, so a more complicated evolutionary history would be necessary to explain the cyathium (Gilbert, 1994); and bisexual flowers are known to occur in ~15 genera in the Euphorbiaceae so it is not implausible that they could occur in *Euphorbia(Prenner and Rudall, 2007)*. Subsequent immunohistology assays showing expression of LEAFY (LFY), a gene associated with floral identity, in cyathium primordia as well as individual floral primordia were interpreted as evidence of the entire cyathium having some degree of floral identity (Prenner et al., 2010). If the cyathium were to be interpreted as a complete flower, the pistillate flower could be interpreted as a pistil and the staminate flowers could be interpreted as stamens. The filiform structures and the involucre could be interpreted as the petals and a connately fused sepal respectively.

Now that a high-quality genome is publicly available for two *Euphorbia* species, the cyathium can be investigated from a comparative molecular developmental perspective. This paper uses transcriptomics and evolutionary analysis to shed more light on the controversial cyathium structure and whether it is homologous to an inflorescence and/or a flower. Transcriptomics of floral development of another species in the Euphorbiaceae, *Jatropha curcas* L., showed that the ABCDE model of floral developmental regulation is mostly conserved in *J. curcas* (Hui et al., 2017). Therefore we can make transcriptomic predictions in *Euphorbia* based on the ABCDE model of floral development.

Under the ABCDE model, different combinations of MADS-box transcription factors combine to specify the four floral whorls: the sepals, petals, stamens, and carpel (Figure 1E). E class

genes are present throughout the flower and help specify all four whorls. C class transcription factors specify the carpel; B and C genes together specify stamens; B and A genes together specify the petal; and A genes alone specify the sepal. D class genes specify ovules. Generally, B and C class genes are conserved between taxa, while A class genes are not highly conserved(Litt and Kramer, 2010). Complete flowers contain all of the floral whorls, while unisexual flowers have either functional stamens or functional carpels but not both. Type I unisexual flowers have sterile carpels/stamens, whereas Type II unisexual flowers, the type posited to be present in the *Euphorbia* cyathium, do not have sterile carpels/stamens. The disappearance of a reproductive whorl in Type II unisexual flowers is hypothesized to result from differential regulation of the MADS-box transcription factors (Feng et al., 2020). Another type of incomplete flowers, achlamydeous flowers, lack a perianth (sepals and petals). Loss of the perianth whorls can result from edits to MADS-box transcription factors or to other transcription factors not related to whorl identity, as in the arabidopsis *petal loss* (*ptl*) mutant(Brewer et al., 2004).

Based on the work of Prenner and Rudall in *Euphorbia spp.* and the ABCDE model, we examine two conflicting hypotheses: the cyathium-as-inflorescence model and the cyathium-as-complete-flower model. Under the cyathium-as-inflorescence model, we expect to see E class genes highly expressed in the pistillate and staminate flowers, potentially expressed in the filiform structures if they constitute reduced flowers, and not expressed in the involucre as it would not be a floral organ. There may or may not be transcriptomic signs of floral organs being initiated that are later aborted (e.g. B class genes expressed in young pistillate flowers and/or perianth-related genes expressed in pistillate or staminate flowers). Under the cyathium-as-complete-flower model, we expect to see expression of E class genes in all of the other “whorls” that would constitute part of the flower including both the filiform structures and involucre which would constitute the “petals” and “sepals”.

MATERIALS AND METHODS

Plant growth conditions—

E. peplus plants were grown under long day conditions with 25 °C day/16 °C night. Seeds were germinated without any pretreatment in closed Phytatrays (P5929; Sigma-Aldrich, Saint Louis, MO, USA) in LM 1-1-1 soil mix, then week-old plants were uppotted to 4-inch pots containing LM 1-1-1 soil mix.

Imaging—

Images of cyathia were taken using a Leica M205 FCA stereo microscope with a DMC6200 camera.

Fixation and dissection for RNA sequencing—

Cyathia from twelve-week-old *E. peplus* plants were collected into ice-cold acetone then stored at -20°C until dissection. Cyathia were dissected under the dissecting scope in acetone on a cold block and dissected organs were put immediately into chilled acetone tubes. 5 biological replicates of 10 different organs were collected, with 10 cyathia dissected for each biological replicate.

Stages used for RNA sequencing—

We used primordia approximately 200 µm tall for our “small primordium” samples. We used cyathium primordia approximately 500 µm tall for our “large primordium” samples. We used staminate flower primordia approximately 200 µm tall for our “young staminate flowers” and pistillate flower primordia approximately 400 µm tall for our “young pistillate flowers”. For the remaining samples, we dissected them from nearly-mature cyathia still enclosed within

unopened cyathophylls when cyathia were approximately 1.5 mm tall not including cyathophylls: nectary, involucre, mature staminate flower, mature pistillate flower, filiform structure. In our organ stages, 'mature' here refers to developmental maturity of form, not reproductive maturity, which roughly coincides with the opening of the cyathophylls.

RNA extraction—

RNA was extracted using the New England Biolabs Monarch Total RNA Miniprep Kit (NEB #T2010, Ipswich, MA, USA). Acetone was removed from the tubes under the dissecting scope, 2-3 beads (BioSpec Products #11079124, Bartlesville, OK, USA) were added, the tubes were placed into liquid nitrogen, and the tissue was ground in a bead beater (Qiagen TissueLyser II, Germantown, MD, USA). 400uL of 1x protection reagent from the kit was added to the ground tissue and beads, and the tissue was ground again in the protection reagent. The liquid was transferred to a new tube and the rest of the kit protocol was followed. RNA was eluted with 100uL nuclease-free water.

RNA purification—

The 100uL of RNA was purified by adding 10uL 3M sodium acetate, 0.75uL GlycoBlue (ThermoFisher #AM9515, Waltham, MA USA) and 275uL ice-cold 100% ethanol, vortexing to mix, precipitating at -80°C overnight, centrifuging at 13000 rpm at 4°C for 30 minutes, washing the pellet twice with 0.5mL ice-cold 75% ethanol, air-drying the pellet for 5 minutes, and resuspending the RNA in 20uL of nuclease-free water.

3' sequencing—

RNA samples were submitted to the Cornell Institute for Biotechnology Genomics Center for 3' mRNA-seq library construction using a QuantSeq FWD kit (Lexogen, Vienna, Austria) followed by 1x75 bp sequencing on an Illumina NextSeq500 (Illumina, Hayward, CA, USA). Sequencing

statistics were calculated with v0.15.0 seqkit stats with parameters -j 10 -T(Shen et al., 2016). (Supplementary File 1).

Orthogroup inference—

To assemble orthogroups, we ran OrthoFinder v2.5.4 (Emms and Kelly, 2019) on the longest isoform protein sequences from *E. peplus*, *E. lathyris*, arabidopsis, soybean, Pará rubber tree, cassava, poplar, and castor bean using the parameters -S diamond -t 40 -a 8(Buchfink et al., 2021). Orthogroups are included in Supplementary File 2.

Identification of ABCDE floral identity genes—

In order to identify floral identity genes of interest, we retrieved the sequences of genes of interest from Uniprot accessed through TAIR and ran blastp with our *E. peplus* proteins file as the query and the parameters -qcov_hsp_perc 80 e-value 1e-10(Camacho et al., 2009). Where BLAST results surfaced multiple candidates with low e-values and high percent identity, a reciprocal BLAST to arabidopsis was used to select the best candidate(s). We further confirmed the results through comparison to the clades identified by OrthoFinder. Gene IDs accessed through TAIR: AT1G69120(AP1)(Bowman et al., 1993); AT4G36920(AP2)(Jofuku et al., 1994); AT3G54340(AP3)(Jack et al., 1992); AT4G18960(AG)(Mizukami and Ma, 1997); AT2G45650(AGL6)(Ma et al., 1991); AT1G69180(CRC)(Alvarez and Smyth, 1999); AT1G65480(FT)(Kobayashi et al., 1999); AT5G60910(AGL8)(Mandel and Yanofsky, 1995); AT5G61850(LFY)(Schultz and Haughn, 1991); AT5G20240(PI)(Goto and Meyerowitz, 1994); AT5G15800(SEP1)(Pelaz et al., 2000); AT1G24260(SEP3)(Pelaz et al., 2000); AT3G58780(SHP1)(Liljegren et al., 2000); AT2G42830(SHP2)(Liljegren et al., 2000); AT2G45660(AGL20)(Lee et al., 2000); AT4G09960(STK)(Favaro et al., 2003); AT5G03840(TFL1)(Kobayashi et al., 1999); AT1G30950(UFO)(Levin and Meyerowitz, 1995).

Transcriptomic analysis—

Raw read files were filtered with fastp v0.23.2 using parameters --cut_right --cut_window_size 5 --cut_mean_quality 20 --length_required 50(Chen et al., 2018). Transcript abundance was generated using *E. peplus* coding sequences with kallisto quant v0.42.4 with parameters --single -b 100 -t 10(Bray et al., 2016). The resulting data was processed in R using the DESeq2 pipeline; all genes with >20 estimated counts across all samples were considered to be expressed and were included in the analysis (Love et al., 2014). Variance-stabilized-transformed data were visualized in heatmaps using the package pheatmap with expression normalized by row and genes clustered by expression pattern(Kolde, n.d.). Of the floral genes of interest investigated (Supplementary File 3), the *E. peplus* homologs for AGL8/FUL (Ep_chr6_g18041), TFL1 (Ep_chr7_g23231) and two paralogs of FT (Ep_chr8_g25489, Ep_chr6_g20489) had expression too low (i.e. 20 or fewer counts across all samples) to be analyzed. Genes are said to be differentially expressed in this paper if padj<0.05. All DESeq filtered untransformed data is included in Supplementary File 4. Comparisons of differentially expressed genes between different tissues are included in Supplementary File 5.

***Transcriptomic comparison between orthologs in arabidopsis and* E. peplus**—

Klepikova floral atlas data for Arabidopsis was downloaded through NCBI PRJNA314076 for the mature sepal, mature anther, mature filament, mature petal, mature carpel, and leaf samples (Klepikova et al., 2015). The raw reads were processed in the fastp/kallisto/DESeq2 pipeline using the same parameters as for the *E. peplus* data. For each orthogroup containing an *E. peplus* gene that was upregulated in the cyathium whorl, we asked whether there were one or more arabidopsis genes in that orthogroup that were upregulated in the arabidopsis flower whorl corresponding to the cyathium whorl.

Biological Process and Molecular Function GO term enrichment analysis—

GO terms were assigned to *E. peplus* genes as specified in Johnson et al 2023 (Johnson et al., 2023); GO terms assigned to each gene are included in Supplementary File 6. In order to find which GO terms were enriched in differentially expressed genes in different organs as compared with cyathophylls, we ran TopGO v.2.46.0 using the “classicCount” option, using Fisher’s test to determine significance(Alexa and Rahnenfuhrer, 2021). TopGO results are included in Supplementary File 7. We summarized and visualized the GO results using REVIGO and downloaded the R scripts to edit the color scale and axis limits (Supek et al., 2011). Terms with a dispensability higher than 0.9 were removed and clustered with other terms for ease of visualization. Only terms with a dispensability less than 0.2 were labeled with text in the visualization.

Manual curation of OrthoFinder clades—

Genes with highly truncated annotations were removed from the analysis (LFY: rubber tree KAF2311631.1, soybean Glyma.04G202002.1.p; PHYA: soybean Glyma.03G227333.1.p; PI: *E. lathyris* Casp05761.t1; SEP1: rubber tree KAF2325472.1). For PI, *E. lathyris* Casp33119.t1 and *E. peplus* Ep_chr8_g25689.t1 were included in the PI phylogenetic analysis despite being classified into their own orthogroup because they clearly were a result of a tandem duplication (Casp33120.1 and Ep_chr8_g25690.t1 were included in the orthogroup with Arabidopsis PI, AT5G20240.1) and because of high reciprocal BLAST scores with Arabidopsis PI.

Gene alignment, gene tree construction and selection analysis—

For each clade of interest (generated by OrthoFinder then manually curated as specified above; see Supplementary File 3 for gene names), a preliminary alignment was generated using MAFFT v7.522 with parameters --maxiterate 5000 --auto --adjustdirectionaccurately --thread 8 --op 3. Then macse v2.07 alignSequences was used to generate a codon-aware alignment with default parameters. For a few alignments, macse generated exclamation points due to

sequencer errors; these were converted to N. Then alignment trimming was performed in trimAl v1.4 with parameter -gappyout. Finally, sequences were exported for phylogeny construction using macse v2.07 exportAlignment -codonForFinalStop NNN -codonForInternalStop NNN -keep_gap_only_sites_ON. To generate phylogenetic trees, RAxML NG v1.2.0 was run with parameters --all --model GTR+G --threads 12 --bs-trees 1000(Kozlov et al., 2019). In order to generate dN/dS values for each branch of the tree, PAML v4.10.6 was run on the best tree from RAxML with model=1 and NSSites=0. In order to test for relaxed selection at sites within genes, PAML v4.10.6 was run on the best tree from RAxML with model=2 and NSSites=2. Alignments were visualized in ESPript 3.0 online (https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi)(Robert and Gouet, 2014). Sites flagged for positive selection based on the Bayes empirical Bayes method (BEB) are included in Supplementary File 8.

Motif analysis—

In order to retrieve a dataset that could be used to detect the canonical LFY binding motif, a bed file of LFY ChIPseq targets (ChIPseq MACS2 q-value 10-10, DANPOS Mnase-seq q-value 10-50) was taken from Supplementary Data 1 from Jin et al 2021(Jin et al., 2021). This was converted to a fasta file of the target sequences using bedtools getfasta(Quinlan and Hall, 2010). Then MEME-ChIP from MEME v5.5.2 was used to detect the LFY canonical binding motif using default parameters except -maxw 25(Machanick and Bailey, 2011). The dLUBS and mLUBS motifs were acquired from the authors of Rieu et al 2023 in pfm format and converted to meme format using jaspar2meme -pfm from MEME v4.11.2(Bailey et al., 2015; Rieu et al., 2023b). Then FIMO from MEME v5.5.2 was run with default parameters on a fasta file of the 3kb upstream of the start of the LFY target genes AP1, AG, AP3, and PI in order to assess whether the motifs were present or not(Grant et al., 2011). MEME FIMO motif search results are included in Supplementary File 9.

Genome assemblies used for analysis—

The Johnson et al *Euphorbia peplus* genome assembly and annotation was accessed through Zenodo(Johnson et al., 2023): https://zenodo.org/records/8033914 The Wang et al. *Euphorbia lathyris* assembly and annotation was accessed through figshare(Wang et al., 2021): https://figshare.com/articles/dataset/High-quality_genome_assembly_of_the_biodiesel_plant_Euphorbia_lathyris/14909913/1 The Liu et al. *Hevea brasiliensis* (rubber tree) assembly and annotation was accessed through NCBI(Liu et al., 2020): https://www.ncbi.nlm.nih.gov/assembly/GCA_010458925.1/. The Xu et al. wild *Ricinus communis* assembly and annotation was accessed through NCBI (Xu et al., 2021): https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_019578655.1/. The Bredeson et al. *Manihot esculenta* v8.1 assembly and annotation was accessed through Phytozome: https://phytozome-next.jgi.doe.gov/info/Mesculenta_v8_1 The arabidopsis Araport 11 genome release was accessed through TAIR(Cheng et al., 2017): https://www.arabidopsis.org/download/index-auto.jsp?dir=%2Fdownload_files%2FGenes%2FAraport11_genome_release%2FAraport11_blastsets. The *Glycine max* (soybean) Wm82.a4.v1 assembly and annotation was accessed through Phytozome(Valliyodan et al., 2019): https://phytozome-next.jgi.doe.gov/info/Gmax_Wm82_a4_v1. The *Populus trichocarpa* v4.1 assembly and annotation was accessed through Phytozome(Tuskan et al., 2006): https://phytozome-next.jgi.doe.gov/info/Ptrichocarpa_v4_1

RESULTS

Analysis of ABCDE gene expression in cyathium components—

We dissected the different organs of the cyathium and sequenced their RNA (Figure 3 A-D). Of 25,473 annotated *E. peplus* genes, 17,164 were expressed in the cyathium. Staminate flowers had the most distinct expression pattern from other organs; a clear separation is visible between the mature staminate flowers, young staminate flowers, and other cyathium organs in the first principal component (PC) of the data (Figure 3E). This data is consistent with many other flowering plant species' transcriptomic pattern of distinct staminate flowers compared with other organs, or distinct stamens compared with other floral organs in complete flowers (Klepikova et al., 2015; Xanthopoulou et al., 2021).

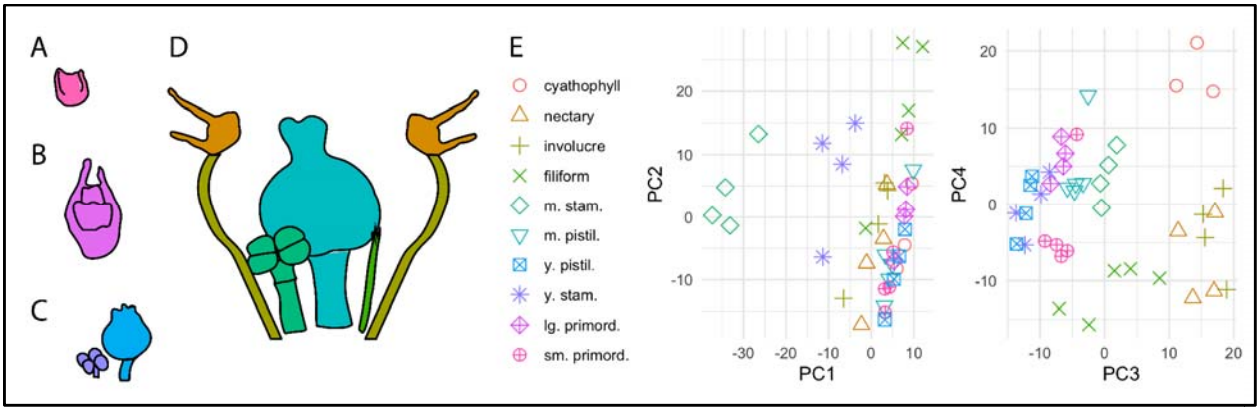


Figure 3

Using BLAST and OrthoFinder, we identified *E. peplus* homologs for all classes of the ABCDE genes and visualized their expression pattern across different organs (Figure 4).

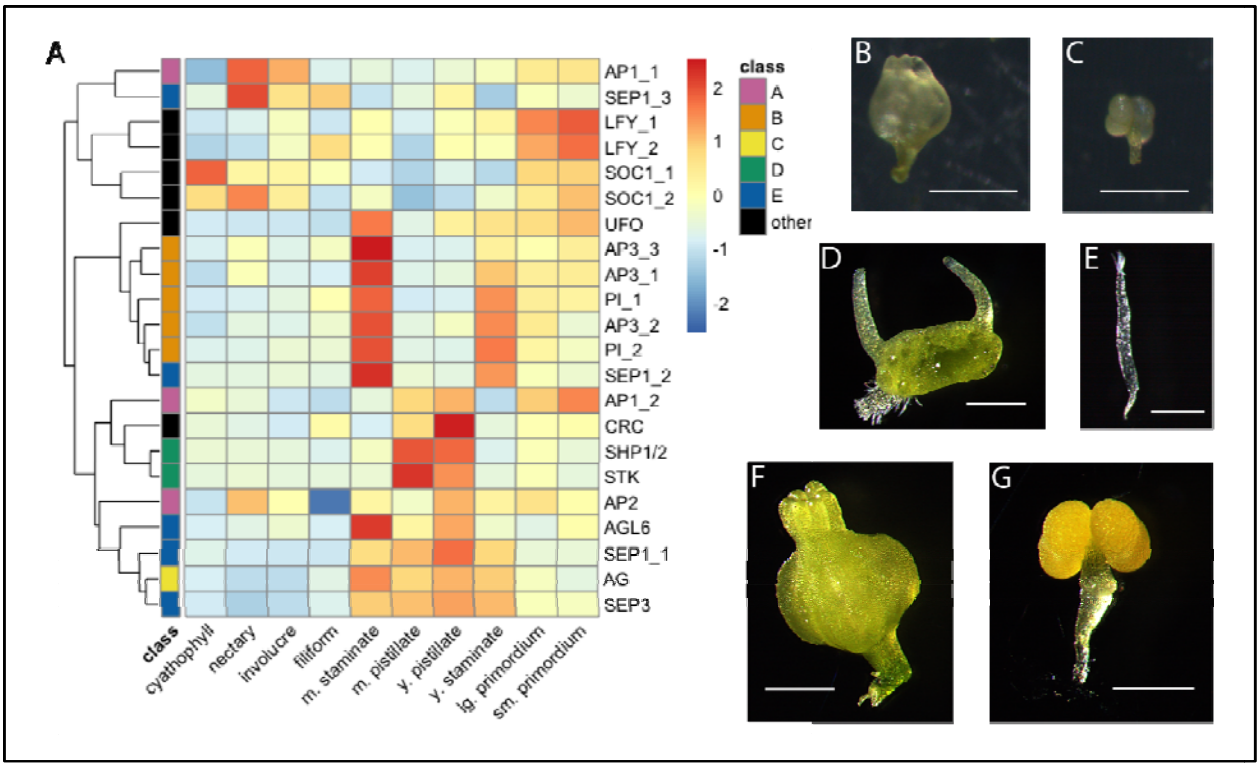


Figure 4

In the ABCDE model, the E class genes SEPALLATA1 (SEP1) and SEPALLATA3 (SEP3) are partially redundant in function and help specify all four floral whorls; a closely related E class gene, AGAMOUS-LIKE 6 (AGL6), also helps specify all of the floral organs and is also involved floral meristem regulation (Dreni and Zhang, 2016; Morel et al., 2019). There are three orthologs of SEP1 in *Euphorbia*, which is consistent with other Euphorbiaceae (Supplementary Figure 1). *E. peplus* SEP1_1 is significantly upregulated in both stages of pistillate flowers compared with cyathophylls, SEP1_2 is significantly upregulated in staminate flowers as compared with pistillate flowers, and SEP 1_3 is significantly upregulated in the filiform structures as compared with cyathophylls. *Euphorbia*'s sole SEP3 ortholog is significantly upregulated in both stages of the pistillate and staminate flowers, both stages of the primordia, and the filiform structures as compared with cyathophylls. AGL6 is not significantly expressed in any tissue as compared with the cyathophylls due to low expression overall.

The B class genes APETALA3 (AP3) and PISTILLATA (PI) specify petal and stamen identity; this role is conserved across bisexual and unisexual flowers (Zhang et al., 2022). All three *E. peplus* AP3 paralogs and both PI paralogs are significantly upregulated in staminate flowers compared with pistillate flowers at both stages studied, in mature staminate flowers as compared with cyathophylls, in young staminate flowers as compared with cyathophylls (except AP3_3), and in filiform structures as compared with cyathophylls. Of the AP3 and PI homologs, only AP3_2 is significantly expressed in the involucres as compared with the cyathophylls.

The C class gene AGAMOUS (AG) is deeply conserved across evolution and helps specify both pistillate and staminate development (Dreni and Kater, 2014). The *E. peplus* ortholog of AG is significantly upregulated in both stages of both pistillate flowers and staminate flowers as compared with cyathophylls, and not significantly upregulated in any other parts of the cyathium.

In arabidopsis, the A class genes APETALA1 (AP1) and APETALA2 (AP2) help specify sepals and help specify petals in conjunction with B class genes, in addition to other functions in floral meristem identity and determinacy; however these genes' functions are not conserved outside the Brassicaceae and are unlikely to be the same in *Euphorbia* as arabidopsis (Litt and Kramer, 2010; Zumajo-Cardona and Pabón-Mora, 2016). There are two homologs of arabidopsis AP1 in *E. peplus*. AP1_1, which is most similar in sequence to arabidopsis AP1 and CAULIFLOWER, is upregulated in the nectary, involucre, filiform structures, both stages of primordia, both stages of staminate flowers, and young pistillate flowers as compared with cyathophylls. *E. peplus* AP1_2 and AP2 are not significantly differentially expressed in any cyathium part as compared with cyathophylls, likely due to low gene expression.

We were also interested in the genes LEAFY (LFY) and UNUSUAL FLORAL ORGANS (UFO) which are known to directly activate many of the ABCDE genes. LFY is a transcription factor that is a key regulator of floral meristem and floral organ identity(Moyroud et al., 2010); it opens up the chromatin around the ABCDE genes, enabling their transcription(Jin et al., 2021). In arabidopsis, UFO interacts with LFY to activate the B class genes AP3 and PI as well as other targets (Rieu et al., 2023b). *E. peplus* has two homologs of LFY: LFY1_1 is significantly upregulated in both stages of primordia as compared with cyathophylls, and LFY_2 is significantly upregulated in small primordia and filiform structures as compared with cyathophylls. *E. peplus* UFO is upregulated in both stages of primordia, both stages of staminate flowers, young pistillate flowers, and in filiform structures, as compared with cyathophylls. It is significantly upregulated in mature staminate flowers as compared with mature pistillate flowers.

The gene CRABS CLAW (CRC), which is necessary for nectary development in both arabidopsis and petunia (Morel et al., 2018; Slavković et al., 2021), is not upregulated in *E. peplus* nectaries as compared with cyathophylls, suggesting an alternative nectary development pathway in *E. peplus*. It is significantly upregulated in both stages of pistillate flowers as well as in filiform structures compared with cyathophylls, and significantly upregulated in both stages of pistillate flowers compared with staminate flowers. The upregulation of CRC in pistillate flowers in *E. peplus* is consistent with CRC's other functions in carpel development and meristem termination and with results from *J. curcas*, where CRC is highly upregulated during developmental stages of pistillate flowers(Xu et al., 2016).

Transcriptomic analysis of differentially expressed genes in cyathium components—Further investigating whether the *E. peplus* cyathium could be homologous to a complete flower, we asked whether the orthogroups containing at least one gene significantly upregulated

in *E. peplus* pistillate flowers as compared with cyathophylls also contained at least one gene significantly upregulated in the arabidopsis carpels compared with leaves; whether orthogroups with *E. peplus* genes upregulated in staminate flowers also contained arabidopsis genes upregulated in stamens; whether orthogroups with *E. peplus* genes upregulated in filiform structures also contained arabidopsis genes upregulated in petals; and whether orthogroups with *E. peplus* genes upregulated in involucres also contained arabidopsis genes upregulated in sepals.

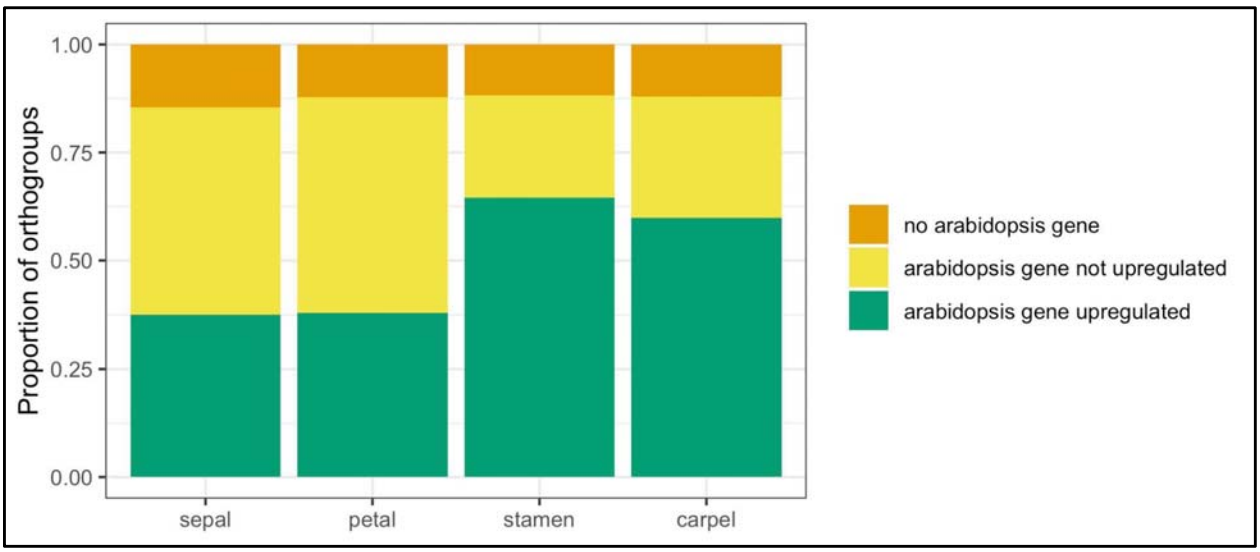


Figure 5

We found that 60% of genes significantly upregulated in *E. peplus* pistillate flowers had an ortholog significantly upregulated in arabidopsis carpels, and 65% of genes significantly upregulated in staminate flowers had an ortholog significantly upregulated in arabidopsis stamens. The filiform structures and involucre both had a much lower level of conservation of expression with arabidopsis floral whorls. 38% of genes significantly upregulated in *E. peplus* filiform structures had an ortholog significantly upregulated in arabidopsis petals, and 38% of genes significantly upregulated in involucres had an ortholog significantly upregulated in arabidopsis sepals.

Because of the lack of shared upregulated orthogroups between involucre and filiform structures with sepal and petal whorls, respectively, and the low level of expression of CRABS CLAW in the nectary, we wanted to investigate what made these parts of the cyathium unique. We performed a gene ontology (GO) term analysis on the genes significantly upregulated in mature staminate flowers, mature pistillate flowers, young staminate flowers, young pistillate flowers, involucres, filiform structures, and nectaries as compared with cyathophylls (Supplementary File 7). We visualized the result for nectaries, involucres, and filiform structures based on semantic similarity of the GO terms (Figure 6).

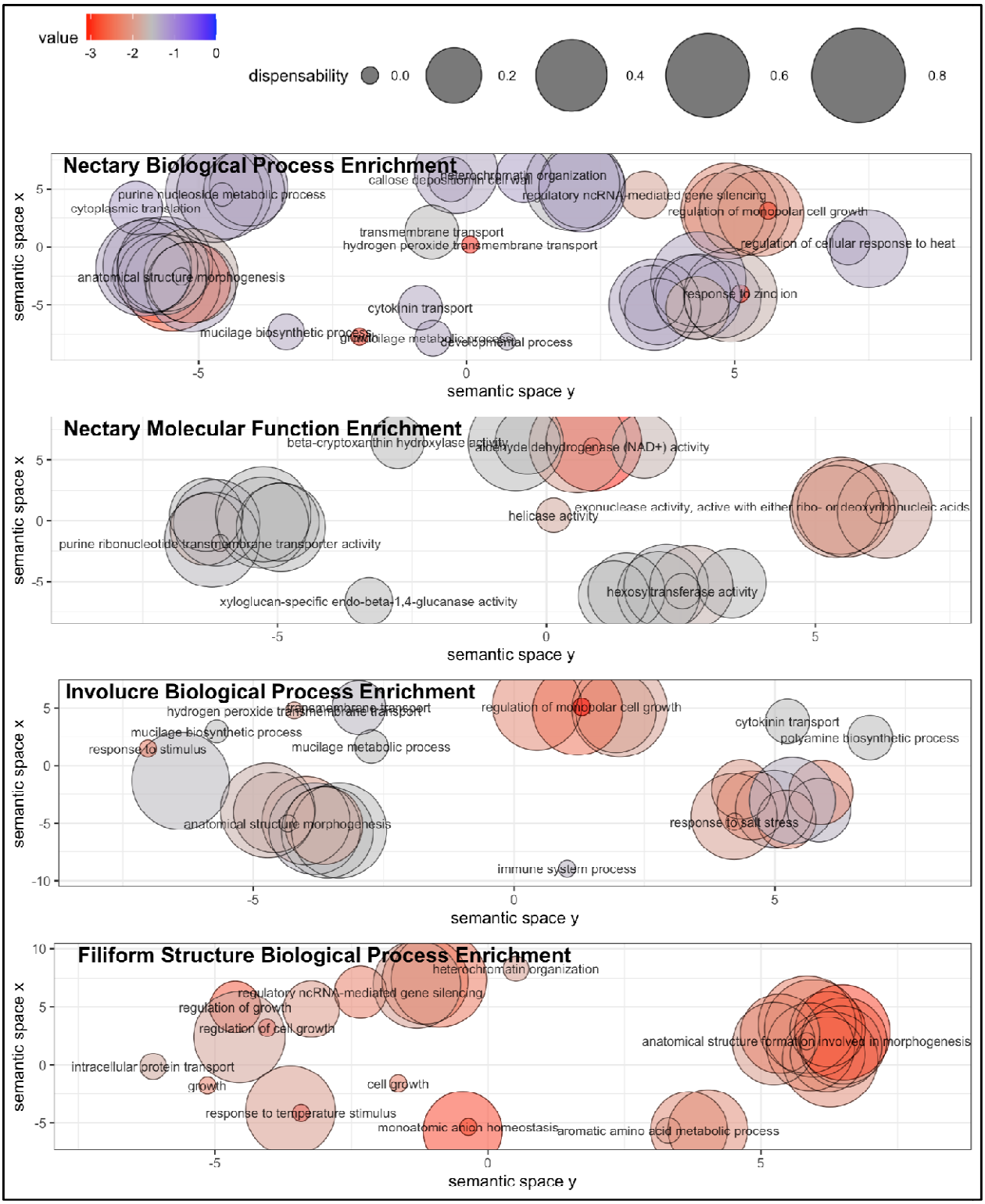

value
-3
-2
-1
0
dispensability
0.0
0.2
0.4
0.6
0.8
Nectary Biological Process Enrichment
callose deposition in cell wall
heterochromatin organization
purine nucleoside metabolic process
cytoplasmic translation
regulatory ncRNA-mediated gene silencing
regulation of monopolar cell growth
transmembrane transport
hydrogen peroxide transmembrane transport
regulation of cellular response to heat
anatomical structure morphogenesis
response to zinc ion
cytokinin transport
mucilage biosynthetic process
mucilage metabolic process
developmental process
semantic space x
semantic space y
Nectary Molecular Function Enrichment
beta-cryptoxanthin hydroxylase activity
aldehyde dehydrogenase (NAD+) activity
helicase activity
exonuclease activity, active with either ribo- or deoxyribonucleic acids
purine ribonucleotide transmembrane transporter activity
hexosyltransferase activity
xyloglucan-specific endo-beta-1,4-glucanase activity
Involucre Biological Process Enrichment
hydrogen peroxide transmembrane transport
regulation of monopolar cell growth
cytokinin transport
mucilage biosynthetic process
polyamine biosynthetic process
response to stimulus
mucilage metabolic process
anatomical structure morphogenesis
response to salt stress
immune system process
Filiform Structure Biological Process Enrichment
heterochromatin organization
regulatory ncRNA-mediated gene silencing
regulation of growth
regulation of cell growth
anatomical structure formation involved in morphogenesis
intracellular protein transport
growth
cell growth
response to temperature stimulus
monoatomic anion homeostasis
aromatic amino acid metabolic process

Figure 6

We found that genes upregulated in the nectary as compared with cyathophylls were significantly enriched for multiple Biological Process GO terms relating to transmembrane transport and specialized metabolism. GO terms relating to gene silencing were also enriched. The Molecular Function GO terms enriched in the nectary included multiple significant terms relating to transmembrane transporter activity and glycosyltransferase activity, similarly to the Biological Process GO terms. ATP and NADH metabolism-related terms were also significantly enriched in the nectaries, suggesting an enrichment in energy-generating redox activities.

Genes upregulated in involucres as compared with cyathophylls were also significantly enriched in genes with Biological Process GO terms relating to specialized metabolite biosynthesis and transmembrane transport. This could potentially indicate that some of the compounds present in the nectaries are at least partially biosynthesized in the involucre tissue connected to the nectaries. The involucre also had several significant GO terms relating to stress, immune responses, and defense responses. Some of these may be related to cytokinin levels; the gene Ep_chr4_g10976, which is upregulated in involucres, is in an orthogroup with AT1G31770(ABCG14), a transmembrane cytokinin transporter. Cytokinin is involved in responses to osmotic stress, salt stress, and other environmental factors but also regulates developmental processes such as inflorescence branching in *J. curcas*(Chen et al., 2019).

The involucre-upregulated gene with the highest p-value, Ep_chr5_g16200, is in an orthogroup with AT5G03680, PETAL LOSS (PTL), which generally represses tissue growth by limiting cell division. (Despite its name, PTL does not regulate petal organ identity, and its expression is not limited to petals– *ptl* mutants' lack of petals may be due to indirect effects of tissue overgrowth disrupting auxin maxima (O'Brien et al., 2015; Takeda et al., 2022).) This gene is upregulated in filiform structures, young pistillate flowers, mature pistillate flowers, nectaries, and both

stages of primordia as compared with cyathophylls, suggesting functions in patterning across multiple cyathium components.

For the genes upregulated in filiform structures as compared with cyathophylls, stamen development Biological Process GO terms were significantly enriched. Heterochromatin formation, gene silencing, and DNA methylation-related terms were also all significantly enriched.

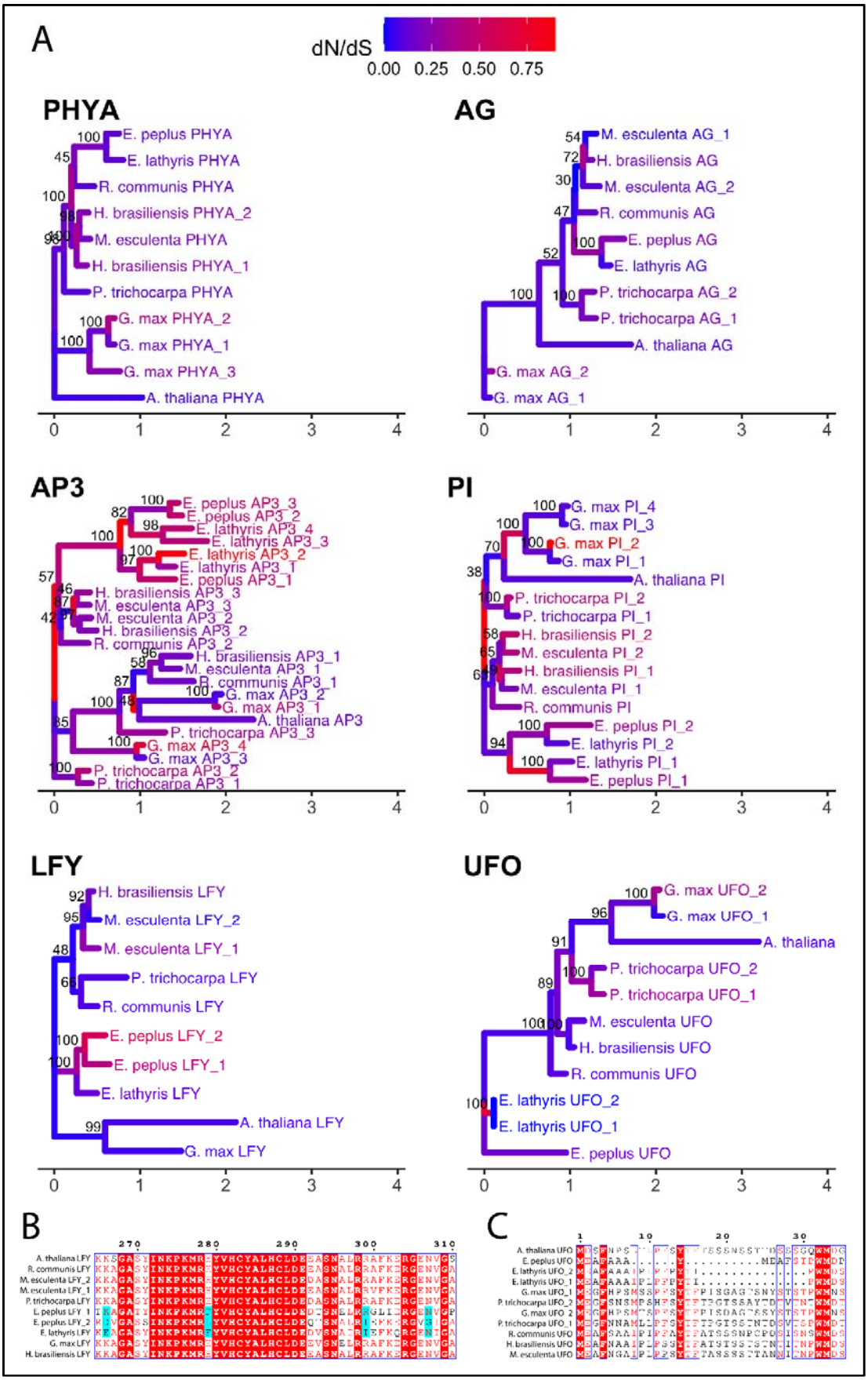
The mature staminate flowers had multiple significant Biological Process GO terms for stamen identity and morphogenesis as well as ATP and NADPH related terms. The young staminate flowers had multiple significant GO terms associated with stamen and pollen development but did not have significant ATP or NADPH related terms. The mature pistillate flowers had two GO terms associated with petal development; the pistillate flower-enriched gene causing this significant GO hit, Ep_chr1_g00250, is in an orthogroup with arabidopsis AT4G16850, KronbladStorleK (KSK; the name is Swedish for petal size) which influences auxin responses during petal development(Miller et al., 2020). Compared with the mature pistillate flower, the younger pistillate flower was enriched for more GO terms relating to reproductive structure development, suggesting that the patterning of the pistillate flower is more active in this earlier stage. Both stages of both the pistillate and staminate flowers had significant GO terms relating to gene silencing; the top five most significant mature pistillate flower terms were all related to gene silencing.

The zinc finger motif-containing gene Ep_chr4_g09148 was in the top 10 most significant upregulated genes in the involucre, nectary, and filiform structures, and was significantly upregulated in every other cyathium tissue studied compared with the involucre. This gene

shares an orthogroup with AT2G33835, FRIGIDA-ESSENTIAL 1 (FES1) which is part of a complex that represses flowering in arabidopsis(Schmitz et al., 2005).

Evolutionary analysis of ABCDE floral identity gene sequences—

Changes in MADS-box binding partners are a driver of floral morphological diversity across many plant families(Bartlett, 2017). For example, duplication and neofunctionalization of SEP genes is involved in the generation of pseudanthia in the Asteraceae(Zhang et al., 2017). Therefore, in addition to analyzing gene expression, we analyzed gene sequence evolution in order to see whether there were changes in *Euphorbia* gene sequence and/or copy number that might reflect changes in binding and function. We performed a codon-aware alignment, tree construction, and dN/dS analysis using the gene sequences from *E. peplus*, *Euphorbia lathyris*, several other Euphorbiaceae species, and the model systems arabidopsis, soybean, and poplar (Figure 7, Supplementary Figure 1, Supplementary Figure 2, Supplementary Figure 3).


A
dN/dS
0.00 0.25 0.50 0.75
PHYA
E. peplus PHYA
E. lathyris PHYA
R. communis PHYA
H. brasiliensis PHYA_2
M. esculenta PHYA
H. brasiliensis PHYA_1
P. trichocarpa PHYA
G. max PHYA_2
G. max PHYA_1
G. max PHYA_3
A. thaliana PHYA
AG
M. esculenta AG_1
H. brasiliensis AG
M. esculenta AG_2
R. communis AG
E. peplus AG
E. lathyris AG
P. trichocarpa AG_2
P. trichocarpa AG_1
A. thaliana AG
G. max AG_2
G. max AG_1
AP3
E. peplus AP3_3
E. peplus AP3_2
E. lathyris AP3_4
E. lathyris AP3_3
E. lathyris AP3_2
E. lathyris AP3_1
E. peplus AP3_1
H. brasiliensis AP3_3
M. esculenta AP3_3
M. esculenta AP3_2
H. brasiliensis AP3_2
R. communis AP3_2
H. brasiliensis AP3_1
M. esculenta AP3_1
R. communis AP3_1
G. max AP3_2
G. max AP3_1
A. thaliana AP3
P. trichocarpa AP3_3
G. max AP3_4
G. max AP3_3
P. trichocarpa AP3_2
P. trichocarpa AP3_1
PI
G. max PI_4
G. max PI_3
G. max PI_2
G. max PI_1
A. thaliana PI
P. trichocarpa PI_2
P. trichocarpa PI_1
H. brasiliensis PI_2
M. esculenta PI_2
H. brasiliensis PI_1
M. esculenta PI_1
R. communis PI
E. peplus PI_2
E. lathyris PI_2
E. lathyris PI_1
E. peplus PI_1
LFY
H. brasiliensis LFY
M. esculenta LFY_2
M. esculenta LFY_1
P. trichocarpa LFY
R. communis LFY
E. peplus LFY_2
E. peplus LFY_1
E. lathyris LFY
A. thaliana LFY
G. max LFY
UFO
G. max UFO_2
G. max UFO_1
A. thaliana
P. trichocarpa UFO_2
P. trichocarpa UFO_1
M. esculenta UFO
H. brasiliensis UFO
R. communis UFO
E. lathyris UFO_2
E. lathyris UFO_1
E. peplus UFO
B
C

497

Figure 7

LFY, UFO, AP3, and PI show high sequence divergence in *Euphorbia* (Figure 7A). Compared with the non-floral gene PHYTOCHROME A (PHYA) and with the C class gene AG, the genes LFY, AP3, and PI have high dN/dS on their *Euphorbia* branches. The *E. peplus* and *E. lathyris* homologs for LFY have an amino acid change from a positive amino acid to a hydrophobic amino acid in a DNA-binding site that is otherwise deeply conserved in vascular plant lineages dating back to the lycophytes (Figure 7B) (Rodríguez-Pelayo et al., 2022). The *E. peplus* and *E. lathyris* homologs for UFO also differ greatly in sequence from other Euphorbiaceae homologs of UFO, and one reason is that the *Euphorbia* versions have deletions near the start of the protein (Figure 7C). Interestingly, many of the AP3 residues flagged as “highly significant” for positive selection in *Euphorbia* are in the K domain in a region involved in tetramer formation(Lai et al., 2019), perhaps indicating a change in binding partner (Supplementary Figure 3).

Because LFY and UFO showed a high level of sequence divergence, in order to investigate whether LFY and UFO binding targets were conserved, we looked for LFY binding motifs and LFY-UFO binding motifs in the putative promoters of homologs of the LFY targets AP3, PI, AG, and AP1 in *E. peplus*, *H. brasiliensis*, and arabidopsis. (AP3 and PI are targets of the LFY-UFO complex as well.) We hypothesized that the binding motifs might not be conserved in *E. peplus*. Both canonical LFY binding motifs and UFO-LFY motifs (mLUBS and dLUBS) are present in the 3kb upstream of these target genes in *E. peplus* (Supplementary File 9), suggesting that at least some of the binding motifs of LFY and LFY-UFO are conserved despite the high divergence in sequence.

DISCUSSION

Across angiosperm evolution, the E class gene SEP3 is generally highly conserved in terms of sequence and expression pattern and usually helps define floral whorls, whereas SEP1 orthologs have been repurposed in different species to regulate bract identity and inflorescence architecture(Dreni and Ferrándiz, 2022). The fact that *Euphorbia*'s SEP3 ortholog is not upregulated in the involucre compared with cyathophylls is inconsistent with the cyathium-as-complete-flower hypothesis.

In the filiform structures, we observed significant expression of B-class and E-class genes, as well as one homolog of LFY, suggesting floral identity. The filiform structures had low levels of RNA overall compared with other structures in the cyathium (Supplementary File 1), which could be explained by the gene ontology analysis showing significant GO terms relating to heterochromatin formation. Our data therefore agrees with the interpretation of the filiform structures as reduced flowers in an inflorescence rather than as petals in a complete flower. Overall, our GO term analysis showed that gene silencing seems to be involved in the creation of multiple cyathium organs.

We did not see transcriptomic signs of abortion of other whorls in the pistillate and staminate flowers (e.g. hypothetically, seeing B class expression in the young pistillate flower could have been an example of that), besides the upregulation of the putative homolog of the arabidopsis perianth gene KSK in the pistillate flower. Organ abortion can happen at many different stages, however, and unisexual flowers do not always show evidence of their missing whorls(Mitchell and Diggle, 2005). The C class gene AGAMOUS is highly expressed in both pistillate flowers and staminate flowers, as would be expected under both of our hypotheses/models.

Euphorbia LFY, UFO, and the B-class genes AP3 and PI show signs of divergence based on phylogenetic distance and relaxed selection. Except for UFO (which may have lacked signs of

relaxed selection due to the gene tree topology), these genes also had multiple sites that were classified as highly significant for positive selection in Bayes Empirical Bayes (BEB) analysis (Supplementary File 8).

LFY and UFO are critically important in specifying floral meristem identity(Rieu et al., 2023a). Many of the originally described loss-of-function or dominant negative *ufo* and *lfy* mutants in arabidopsis had an inflorescence subtended by a bract in place of a flower(Weigel et al., 1992; Levin and Meyerowitz, 1995). UFO mutants in other species also have an inflorescence in place of a flower, e.g. *anantha* mutants in tomato(Lippman et al., 2008). In the Euphorbiaceae, the LFY homolog in *J. curcas* promotes the transition from inflorescence to floral meristem identity. Overexpression of JcLFY causes early flowering, and its co-suppression delays flower formation, causes more inflorescence branches, and causes sepal-like organs(Tang et al., 2016). As compared with its relatives such as *Anthostema* (Figure 2), the *Euphorbia* inflorescence is very compact structurally. Cyathia have a reduced axis, different placement of their flowers, and an earlier emergence of pistillate flower primordia compared with staminate flower primordia than inflorescences of related genera (Gagliardi et al., 2018). These differences could not be explained by differential elongation of organs. This suggests that a difference in timing or spatial location of the specification and/or termination of the floral meristems led to the cyathium structure. The drastic changes in sequence of LFY and UFO in *Euphorbia* could contribute to those functional changes in floral meristem specification.

One interesting side note on the expression of UFO is that in *J. curcas*, UFO is downregulated in both mature staminate flowers and mature pistillate flowers compared to younger stages(Chen et al., 2016), whereas in our *E. peplus* dataset we saw the highest expression of UFO in mature staminate flowers. This either suggests a repurposing of UFO, or that developmentally mature but not reproductively mature (i.e. not elongated) staminate flowers

aren't yet mature transcriptomically compared with pistillate flowers. As UFO functions in the development of individual perianth organs as well as in the broader activation of the B class genes(Takeda et al., 2022), some of the changes in UFO expression and sequence in *Euphorbia* could be related to the suppression of perianth development in addition to potential functions in defining the floral meristem.

Genes in the AP1 clade can also be involved in floral meristem specification– and are also highly diverged in *Euphorbia*. In arabidopsis, AGL79 interacts with SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1) to regulate the expression of TERMINAL FLOWER 1 (TFL1) which controls inflorescence meristem identity (Yang et al., 2023). *Euphorbia* appears to be "missing" its AP1-CAL homolog in the Euphorbiaceae AGL79 clade in our gene tree because that gene, AP1_2, is so diverged that it appears elsewhere in the tree (Supplementary Figure 2).

Based on our results, we hypothesize that the cyathium is indeed an inflorescence, not a complete flower, and that there are two key genetic innovations that led to the cyathium: 1. Alteration of LFY function to change the timing of floral meristem specification and floral development to create a compact inflorescence; and 2. Changing the transcription factor binding partners of AP3 and PI so that they are no longer involved in perianth formation, deleting the structure of the perianth.

The modification of AP3 and PI to change their binding partners and function in order to delete the perianth would constitute a destruction of the system where overlapping ABC genes create four whorls of a flower; in that case, Euphorbia would be similar to gymnosperms where B and C class expression together specifies pollenate organs and C class alone specifies ovulate organs but there is no perianth (Figure 8)(Irish, 2017). Hypothetically, because the B class

proteins would no longer be required to work together with A class members, the B class genes would be freed to take on new functions, which could help explain some of the high diversity in cyathium form in *Euphorbia*.

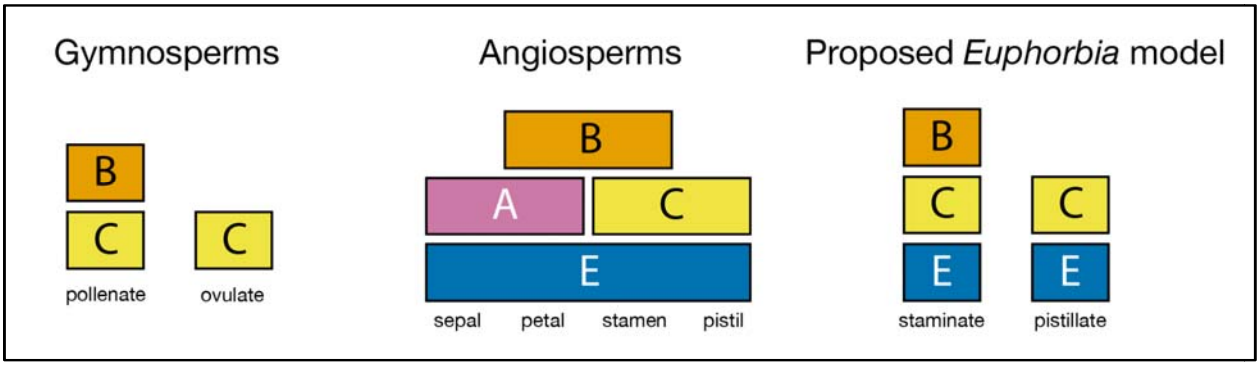


Figure 8

Our hypothesis could be consistent with Prenner and Rudall's interpretation of the Australian genus *Calycopeplus* and the New Caledonia tree *Neoguillauminia cleopatra* as "missing links" towards the *Euphorbia* cyathium. Because *Calycopeplus* and *Neoguillauminia* have a similarly compact inflorescence structure to that of *Euphorbia* (compared with e.g. *Anthostema*) but still have perianths, we hypothesize that a step-wise evolutionary process occurred where an inflorescence-disrupting event (i.e. disruption of LFY function) happened in the common ancestor of *Calycopeplus*, *Neoguillauminia,* and *Euphorbia* but the alteration of B class genes to delete the perianth happened only in *Euphorbia*. Future studies could sequence the floral genes of *Calycopeplus*, *Neoguillauminia*, and *Anthostema* to test this hypothesis.

In addition to expanding our analyses to more taxa, future studies could test our model through experimental investigations of the function of the UFO-LFY complex and its DNA binding targets in *Euphorbia*, e.g. through electrophoretic mobility shift assays (EMSA), microscale thermophoresis (MST), or DAP-seq. Additionally, yeast two-hybrid or similar techniques could

be used to check the transcription factor binding partners of *Euphorbia*'s B-class genes. It would also be interesting to use ATAC-seq or another method to determine chromatin accessibility across the genome and speculate on what might be being silenced in the different cyathium structures. *In situ* hybridization, if optimized in this system, could also potentially reveal more spatially and temporally resolved patterns of gene expression. We are excited to see what future work reveals about the developmental genetics involved in creating the complex structure of the cyathium.

ACKNOWLEDGEMENTS

Thank you to Jacob Landis for extremely helpful advice on running PAML. A warm thank you to François Parcy, Jérémy Lucas and Laura Turchi for sending the motif files for mLUBS and dLUBS. Thank you to the Cornell Weill growth chamber plant maintenance staff, especially Corbin Lay, for their excellent upkeep of the plants used. Thank you to the Frank Lab for helpful comments on the manuscript. This work was funded by a Cornell Institute of Biotechnology Seed Grant, National Science Foundation Integrative Organismal Systems (NSF IOS) award 1942437, and startup funds from the College of Agriculture and Life Sciences at Cornell University to Margaret Frank.

AUTHOR CONTRIBUTIONS

A.R.J., A.B., and M.H.F. all conceived of the study. Under the supervision of A.R.J., A.B. collected the images of cyathia and dissected cyathia components, and performed preliminary dissections and RNA extractions that contributed to the experimental design for the RNAseq

experiment. A.R.J. performed the dissection for the RNAseq experiment. A.B. and A.R.J. both contributed to the RNA extraction for the RNAseq experiment. A.B. and A.R.J. collaborated on the initial sequence alignment and phylogeny construction to assess conservation of sequence similarity in floral genes. A.R.J. performed all other data analysis. A.R.J. wrote the first draft of the manuscript, after which all authors contributed to reviewing and editing the manuscript.

DATA AVAILABILITY

The raw RNAseq data generated in this study are available from NCBI GenBank under the accession PRJNA1063728. All code used in the data analysis is available on Github at https://github.com/ariellerjohnson/Cyathium-2024.

LITERATURE CITED

Alexa, A., and J. Rahnenfuhrer. 2021. topGO: enrichment analysis for gene ontology. *R package version* 2.46: 2021.

Alvarez, J., and D. R. Smyth. 1999. CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. *Development* 126: 2377–2386.

Bailey, T. L., J. Johnson, C. E. Grant, and W. S. Noble. 2015. The MEME Suite. *Nucleic acids research* 43: W39–49.

Bartlett, M. E. 2017. Changing MADS-Box Transcription Factor Protein-Protein Interactions as a Mechanism for Generating Floral Morphological Diversity. *Integrative and comparative*

biology 57: 1312–1321.

Bowman, J. L., J. Alvarez, D. Weigel, E. M. Meyerowitz, and D. R. Smyth. 1993. Control of flower development in Arabidopsis thaliana by APETALA1 and interacting genes. *Development* 119: 721–743.

Bray, N. L., H. Pimentel, P. Melsted, and L. Pachter. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology* 34: 525–527.

Brewer, P. B., P. A. Howles, K. Dorian, M. E. Griffith, T. Ishida, R. N. Kaplan-Levy, A. Kilinc, and D. R. Smyth. 2004. PETAL LOSS, a trihelix transcription factor gene, regulates perianth architecture in the Arabidopsis flower. *Development* 131: 4035–4045.

Bruyns, P. V. 2022. Euphorbia in Southern Africa: Volume 1. Springer Nature.

Buchfink, B., K. Reuter, and H.-G. Drost. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature methods* 18: 366–368.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.

Cheng, C.-Y., V. Krishnakumar, A. P. Chan, F. Thibaud-Nissen, S. Schobel, and C. D. Town. 2017. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *The Plant journal: for cell and molecular biology* 89: 789–804.

Chen, M.-S., B.-Z. Pan, Q. Fu, Y.-B. Tao, J. Martínez-Herrera, L. Niu, J. Ni, et al. 2016. Comparative Transcriptome Analysis between Gynoecious and Monoecious Plants Identifies Regulatory Networks Controlling Sex Determination in Jatropha curcas. *Frontiers in plant science* 7: 1953.

Chen, M.-S., M.-L. Zhao, G.-J. Wang, H.-Y. He, X. Bai, B.-Z. Pan, Q.-T. Fu, et al. 2019.

Transcriptome analysis of two inflorescence branching mutants reveals cytokinin is an important regulator in controlling inflorescence architecture in the woody plant Jatropha curcas. *BMC plant biology* 19: 468.

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

Claßen-Bockhoff, R., and H. Frankenhäuser. 2020. The 'Male Flower' of Ricinus communis (Euphorbiaceae) Interpreted as a Multi-Flowered Unit. *Frontiers in cell and developmental biology* 8: 313.

Czechowski, T., E. Forestier, S. H. Swamidatta, A. D. Gilday, A. Cording, T. R. Larson, D. Harvey, et al. 2022. Gene discovery and virus-induced gene silencing reveal branched pathways to major classes of bioactive diterpenoids in Euphorbia peplus. *Proceedings of the National Academy of Sciences of the United States of America* 119: e2203890119.

Dreni, L., and C. Ferrándiz. 2022. Tracing the Evolution of the SEPALLATA Subfamily across Angiosperms Associated with Neo- and Sub-Functionalization for Reproductive and Agronomically Relevant Traits. *Plants* 11.

Dreni, L., and M. M. Kater. 2014. MADS reloaded: evolution of the AGAMOUS subfamily genes. *The New phytologist* 201: 717–732.

Dreni, L., and D. Zhang. 2016. Flower development: the evolutionary history and functions of the AGL6 subfamily MADS-box genes. *Journal of experimental botany* 67: 1625–1638.

Dressler, S., M. Schmidt, and G. Zizka. 2014. Introducing African plants—A photo guide—an interactive photo data□base and rapid identification tool for continental Africa. *Taxon* 63: 1159–1161.

Emms, D. M., and S. Kelly. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome biology* 20: 238.

Favaro, R., A. Pinyopich, R. Battaglia, M. Kooiker, L. Borghi, G. Ditta, M. F. Yanofsky, et al. 2003. MADS-box protein complexes control carpel and ovule development in Arabidopsis. *The Plant cell* 15: 2603–2611.

Feng, G., B. J. Sanderson, K. Keefover-Ring, J. Liu, T. Ma, T. Yin, L. B. Smart, et al. 2020. Pathways to sex determination in plants: how many roads lead to Rome? *Current opinion in plant biology* 54: 61–68.

Gagliardi, K. B., I. Cordeiro, and D. Demarco. 2018. Structure and development of flowers and inflorescences in Peraceae and Euphorbiaceae and the evolution of pseudanthia in Malpighiales. *PloS one* 13: e0203954.

Gilbert, M. G. 1994. The relationships of the Euphorbieae (Euphorbiaceae). *Annals of the Missouri Botanical Garden. Missouri Botanical Garden* 81: 283.

Goto, K., and E. M. Meyerowitz. 1994. Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. *Genes & development* 8: 1548–1560.

Grant, C. E., T. L. Bailey, and W. S. Noble. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27: 1017–1018.

Hui, W., Y. Yang, G. Wu, C. Peng, X. Chen, and M. Z. Zayed. 2017. Transcriptome profile analysis reveals the regulation mechanism of floral sex differentiation in Jatropha curcas L. *Scientific reports* 7: 16421.

Irish, V. 2017. The ABC model of floral development. *Current biology: CB* 27: R887–R890.

Jack, T., L. L. Brockman, and E. M. Meyerowitz. 1992. The homeotic gene APETALA3 of

Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. *Cell* 68: 683–697.

Jin, R., S. Klasfeld, Y. Zhu, M. Fernandez Garcia, J. Xiao, S.-K. Han, A. Konkol, and D. Wagner. 2021. LEAFY is a pioneer transcription factor and licenses cell reprogramming to floral fate. *Nature communications* 12: 626.

Jofuku, K. D., B. G. den Boer, M. Van Montagu, and J. K. Okamuro. 1994. Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *The Plant cell* 6: 1211–1225.

Johnson, A. R., Y. Yue, S. B. Carey, S. J. Park, L. H. Kruse, A. Bao, A. Pasha, et al. 2023. Chromosome-level genome assembly of Euphorbia peplus, a model system for plant latex, reveals that relative lack of Ty3 transposons contributed to its small genome size. *Genome biology and evolution* 15: evad018.

Kaplan, D., and C. D. Specht. 2022. Kaplan's Principles of Plant Morphology. CRC Press.

Klepikova, A. V., M. D. Logacheva, S. E. Dmitriev, and A. A. Penin. 2015. RNA-seq analysis of an apical meristem time series reveals a critical point in Arabidopsis thaliana flower initiation. *BMC genomics* 16: 466.

Kobayashi, Y., H. Kaya, K. Goto, M. Iwabuchi, and T. Araki. 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science* 286: 1960–1962.

Kolde. Pheatmap: pretty heatmaps. *R package version*.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35: 4453–4455.

Lai, X., H. Daher, A. Galien, V. Hugouvieux, and C. Zubieta. 2019. Structural Basis for Plant MADS Transcription Factor Oligomerization. *Computational and structural biotechnology journal* 17: 946–953.

Lee, H., S. S. Suh, E. Park, E. Cho, J. H. Ahn, S. G. Kim, J. S. Lee, et al. 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes & development* 14: 2366–2376.

Levin, J. Z., and E. M. Meyerowitz. 1995. UFO: an Arabidopsis gene involved in both floral meristem and floral organ development. *The Plant cell* 7: 529–548.

Li, C., J. Su, X. Liu, S. Chen, and L. He. 2014. Pistillate Flower Development and Stigma Receptivity of Euphorbia pulcherrima. *Agricultural Science & Technology* 15: 1671–1675.

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

Lippman, Z. B., O. Cohen, J. P. Alvarez, M. Abu-Abied, I. Pekker, I. Paran, Y. Eshed, and D. Zamir. 2008. The making of a compound inflorescence in tomato and related nightshades. *PLoS biology* 6: e288.

Litt, A., and E. M. Kramer. 2010. The ABC model and the diversification of floral organ identity. *Seminars in cell & developmental biology* 21: 129–137.

Liu, J., C. Shi, C.-C. Shi, W. Li, Q.-J. Zhang, Y. Zhang, K. Li, et al. 2020. The Chromosome-Based Rubber Tree Genome Provides New Insights into Spurge Genome Evolution and Rubber Biosynthesis. *Molecular plant* 13: 336–350.

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and

dispersion for RNA-seq data with DESeq2. *Genome biology* 15: 550.

Machanick, P., and T. L. Bailey. 2011. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* 27: 1696–1697.

Ma, H., M. F. Yanofsky, and E. M. Meyerowitz. 1991. AGL1-AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. *Genes & development* 5: 484–495.

Mandel, M. A., and M. F. Yanofsky. 1995. The Arabidopsis AGL8 MADS box gene is expressed in inflorescence meristems and is negatively regulated by APETALA1. *The Plant cell* 7: 1763–1771.

Miller, C. N., J. Dumenil, F. H. Lu, C. Smith, N. McKenzie, V. Chapman, J. Ball, et al. 2020. Variation in the expression of a transmembrane protein influences cell growth in Arabidopsis thaliana petals by altering auxin responses. *BMC plant biology* 20: 482.

Mitchell, C. H., and P. K. Diggle. 2005. The evolution of unisexual flowers: morphological and functional convergence results from diverse developmental transitions. *American journal of botany* 92: 1068–1076.

Mizukami, Y., and H. Ma. 1997. Determination of Arabidopsis floral meristem identity by AGAMOUS. *The Plant cell* 9: 393–408.

Morel, P., P. Chambrier, V. Boltz, S. Chamot, F. Rozier, S. Rodrigues Bento, C. Trehin, et al. 2019. Divergent Functional Diversification Patterns in the SEP/AGL6/AP1 MADS-Box Transcription Factor Superclade. *The Plant cell* 31: 3033–3056.

Morel, P., K. Heijmans, K. Ament, M. Chopy, C. Trehin, P. Chambrier, S. Rodrigues Bento, et al. 2018. The Floral C-Lineage Genes Trigger Nectary Development in Petunia and

Arabidopsis. *Plant Cell* 30: 2020–2037.

Moyroud, E., E. Kusters, M. Monniaux, R. Koes, and F. Parcy. 2010. LEAFY blossoms. *Trends in plant science* 15: 346–352.

O'Brien, M., R. N. Kaplan-Levy, T. Quon, P. G. Sappl, and D. R. Smyth. 2015. PETAL LOSS, a trihelix transcription factor that represses growth in Arabidopsis thaliana, binds the energy-sensing SnRK1 kinase AKIN10. *Journal of experimental botany* 66: 2475–2485.

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

Prenner, G., N. I. Cacho, D. Baum, and P. J. Rudall. 2010. Is LEAFY a useful marker gene for the flower–inflorescence boundary in the Euphorbia cyathium? *Journal of experimental botany* 62: 345–350.

Prenner, G., S. D. Hopper, and P. J. Rudall. 2008. Pseudanthium development in Calycopeplus paucifolius, with particular reference to the evolution of the cyathium in Euphorbieae (Euphorbiaceae–Malpighiales). *Australian systematic botany* 21: 153–161.

Prenner, G., and P. J. Rudall. 2007. Comparative ontogeny of the cyathium in Euphorbia (Euphorbiaceae) and its allies: exploring the organ-flower-inflorescence boundary. *American journal of botany* 94: 1612–1629.

Quinlan, A. R., and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842.

Rieu, P., M. Arnoux-Courseaux, G. Tichtinsky, and F. Parcy. 2023a. Thinking outside the F-box: how UFO controls angiosperm development. *The New phytologist* 240: 945–959.

Rieu, P., L. Turchi, E. Thévenon, E. Zarkadas, M. Nanao, H. Chahtane, G. Tichtinsky, et al.

2023b. The F-box protein UFO controls flower development by redirecting the master transcription factor LEAFY to new cis-elements. *Nature plants* 9: 315–329.

Robert, X., and P. Gouet. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research* 42: W320–4.

Rodríguez-Pelayo, C., B. A. Ambrose, A. Vasco, J. F. Alzate, and N. Pabón-Mora. 2022. Evolution and expression of LEAFY genes in ferns and lycophytes. *EvoDevo* 13: 2.

Schmitz, R. J., L. Hong, S. Michaels, and R. M. Amasino. 2005. FRIGIDA-ESSENTIAL 1 interacts genetically with FRIGIDA and FRIGIDA-LIKE 1 to promote the winter-annual habit of Arabidopsis thaliana. *Development* 132: 5471–5478.

Schultz, E. A., and G. W. Haughn. 1991. LEAFY, a Homeotic Gene That Regulates Inflorescence Development in Arabidopsis. *The Plant cell* 3: 771–781.

Shen, W., S. Le, Y. Li, and F. Hu. 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PloS one* 11: e0163962.

Slavković, F., C. Dogimont, H. Morin, A. Boualem, and A. Bendahmane. 2021. The Genetic Control of Nectary Development. *Trends in plant science* 26: 260–271.

Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS one* 6: e21800.

Takeda, S., Y. Hamamura, T. Sakamoto, S. Kimura, M. Aida, and T. Higashiyama. 2022. Non-cell-autonomous regulation of petal initiation in Arabidopsis thaliana. *Development* 149.

Tang, M., Y.-B. Tao, Q. Fu, Y. Song, L. Niu, and Z.-F. Xu. 2016. An ortholog of LEAFY in Jatropha curcas regulates flowering time and floral organ development. *Scientific reports* 6: 37306.

Tuskan, G. A., S. Difazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, et al. 2006. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* 313: 1596–1604.

Valliyodan, B., S. B. Cannon, P. E. Bayer, S. Shu, A. V. Brown, L. Ren, J. Jenkins, et al. 2019. Construction and comparison of three reference-quality genome assemblies for soybean. *The Plant journal: for cell and molecular biology* 100: 1066–1082.

Wang, M., Z. Gu, Z. Fu, and D. Jiang. 2021. High-quality genome assembly of an important biodiesel plant, Euphorbia lathyris L. *DNA research: an international journal for rapid publication of reports on genes and genomes* 28.

Weigel, D., J. Alvarez, D. R. Smyth, M. F. Yanofsky, and E. M. Meyerowitz. 1992. LEAFY controls floral meristem identity in Arabidopsis. *Cell* 69: 843–859.

Western Australian Herbarium, Biodiversity, Conservation Science, Department of Biodiversity, Conservation, and Attractions. Calycopeplus Planch. Website https://florabase.dbca.wa.gov.au/browse/profile/21684 [accessed 25 January 2024].

Xanthopoulou, A., J. Montero-Pau, B. Picó, P. Boumpas, E. Tsaliki, H. S. Paris, A. Tsaftaris, et al. 2021. A comprehensive RNA-Seq-based gene expression atlas of the summer squash (Cucurbita pepo) provides insights into fruit morphology and ripening mechanisms. *BMC genomics* 22: 341.

Xu, G., J. Huang, Y. Yang, and Y.-A. Yao. 2016. Transcriptome Analysis of Flower Sex Differentiation in Jatropha curcas L. Using RNA Sequencing. *PloS one* 11: e0145613.

Xu, W., D. Wu, T. Yang, C. Sun, Z. Wang, B. Han, S. Wu, et al. 2021. Genomic insights into the origin, domestication and genetic basis of agronomic traits of castor bean. *Genome biology* 22: 113.

Yang, H., P. Zhang, D. Guo, N. Wang, H. Lin, X. Wang, and L. Niu. 2023. Transcriptional repressor AGL79 positively regulates flowering time in Arabidopsis. *Journal of plant physiology* 285: 153985.

Zhang, X., L. Pan, W. Guo, Y. Li, and W. Wang. 2022. A convergent mechanism of sex determination in dioecious plants: Distinct sex-determining genes display converged regulation on floral B-class genes. *Frontiers in plant science* 13: 953445.

Zumajo-Cardona, C., and N. Pabón-Mora. 2016. Evolution of the APETALA2 Gene Lineage in Seed Plants. *Molecular biology and evolution* 33: 1818–1832.

FIGURE LEGENDS

Figure 1: The *Euphorbia* cyathium is a unique reproductive structure. A. Diagram of a *Euphorbia peplus* cyathium. B. *E. peplus* cyathium and adjacent primordia and cyathophylls. Scale=2mm. C. *E. peplus* cyathium and adjacent primordia with adjacent cyathophylls removed. Scale=1mm. D. *E. peplus* cyathium with involucre and nectaries removed. Scale=500 microns. E. Diagram of the ABCDE model of floral development and the four whorls of a traditional flower; D class genes, which control ovule development, are not shown.

Figure 2: Close relatives of *Euphorbia*, *Calycopeplus* and *Neoguillauminia*, have perianths around their pistillate and staminate flowers and a condensed inflorescence similar to that of *Euphorbia*. *Anthostema*, which is less closely related, has perianths around its pistillate and staminate flowers and a less-condensed inflorescence.

Figure 3: Staminate flowers have a distinct expression pattern. A-D. Illustration of cyathium parts used for RNAseq sampling. A. Small primordium. B. Large primordium. C. Young staminate flower and young pistillate flower. D. Mature organs: nectary, involucre, staminate flower, pistillate flower, filiform structure, involucre, nectary. (Duplicates of organs are omitted for simplicity. Cyathophyll is not shown.) E. PCA plot of variance-stabilized-transformed data for the 500 genes with the most variance. PC1 explains 22.4% of the variance, PC2 explains 18.9% of the variance, PC3 explains 16.7% of the variance, and PC4 explains 10.6% of the variance.

Figure 4: E class gene SEP3 is highly expressed in the pistillate and staminate flowers. A. Heatmap shows expression of ABCDE homologs and other floral identity genes in different organs of the *E. peplus* cyathium. B-N. Images of the organs sampled. B. Young pistillate flower. C. Young staminate flower. D. Nectary. E. Filiform structure. F. Mature pistillate flower. G. Mature staminate flower. Scale bars = 250 microns.

Figure 5: *E. peplus* staminate flowers and pistillate flowers have higher transcriptomic conservation with arabidopsis stamens and carpels than *E. peplus* filiform structures and involucres have with arabidopsis petals and sepals. Total number of orthogroups with upregulated genes in each *E. peplus* structure (compared with *E. peplus* cyathophylls) shown in plot: involucre (vs. sepal): 649; filiform structure (vs. petal): 1422; staminate flower (vs. stamen): 1424; pistillate flower (vs. carpel): 473.

Figure 6: Biological Process GO terms for genes significantly upregulated in the nectary, involucre, and filiform structures as compared with the cyathophyll and Molecular Function GO terms for genes significantly upregulated in the nectary as compared with the cyathophyll are summarized using REVIGO. “Value” is the log10(p-value); a lower value, towards the red side

of the color scale, is more significant. The size of the points, the dispensability, represents the semantic similarity to other terms in the dataset.

Figure 7: LFY, UFO, and B class genes show signs of divergence and relaxed selection in *Euphorbia*. A. dN/dS values shown on gene trees of floral genes of interest; RAxML support value from 1000 bootstraps is shown at nodes. PHYA is shown as a non-floral “control”. B. Visualization of amino acid alignment in a highly conserved region of LFY; site 279 is a DNA binding site that is highly conserved (except in *Euphorbia*). Cyan color shows sites in the LFY protein alignment rated as “highly significant” for positive selection in the *Euphorbia* clade by PAML. C. Visualization of amino acid alignment at the start of the UFO protein showing deletions in *Euphorbia* (“Ep” and “Casp” genes). Gene names and their gene ID equivalents are listed in Supplementary File 3.

Figure 8: Proposed new model of *Euphorbia* cyathium development in relation to angiosperm and gymnosperm reproductive development.