



OPEN ACCESS

EDITED BY

Sean W. Graham,
University of British Columbia, Canada

REVIEWED BY

Alex Twyford,
The University of Edinburgh,
United Kingdom
Jocelyn Hall,
University of Alberta, Canada

*CORRESPONDENCE

Chelsea D. Specht
cdspecht@cornell.edu

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Plant Systematics and Evolution,
a section of the journal
Frontiers in Plant Science

RECEIVED 12 February 2022

ACCEPTED 27 June 2022

PUBLISHED 08 September 2022

CITATION

Valderrama E, Landis JB, Skinner D,
Maas PJM, Maas-van de Kramer H,
André T, Grunder N, Sass C,
Pinilla-Vargas M, Guan CJ, Phillips HR,
Almeida AMRd and Specht CD (2022)
The genetic mechanisms underlying
the convergent evolution of pollination
syndromes in the Neotropical radiation
of *Costus* L.
Front. Plant Sci. 13:874322.
doi: 10.3389/fpls.2022.874322

COPYRIGHT

© 2022 Valderrama, Landis, Skinner,
Maas, Maas-van de Kramer, André,
Grunder, Sass, Pinilla-Vargas, Guan,
Phillips, Almeida and Specht. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

The genetic mechanisms underlying the convergent evolution of pollination syndromes in the Neotropical radiation of *Costus* L.

Eugenio Valderrama^{1†}, Jacob B. Landis^{1,2†}, Dave Skinner³,
Paul J. M. Maas⁴, Hiltje Maas-van de Kramer⁴, Thiago André⁵,
Nikolaus Grunder⁶, Chodon Sass⁷, Maria Pinilla-Vargas¹,
Clarice J. Guan¹, Heather R. Phillips¹,
Ana Maria Rocha de Almeida⁶ and Chelsea D. Specht^{1*}

¹School of Integrative Plant Science, Section of Plant Biology and the L.H. Bailey Hortorium, Cornell University, Ithaca, NY, United States, ²BTI Computational Biology Center, Boyce Thompson Institute, Ithaca, NY, United States, ³Le Jardin Ombragé, Tallahassee, FL, United States, ⁴Section Botany, Naturalis Biodiversity Center, Leiden, Netherlands, ⁵Departamento de Botânica, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, DF, Brazil, ⁶Department of Biological Sciences, California State University, East Bay, Hayward, CA, United States, ⁷University and Jepson Herbaria, University of California, Berkeley, Berkeley, CA, United States

Selection together with variation in floral traits can act to mold floral form, often driven by a plant's predominant or most effective pollinators. To investigate the evolution of traits associated with pollination, we developed a phylogenetic framework for evaluating tempo and mode of pollination shifts across the genus *Costus* L., known for its evolutionary toggle between traits related to bee and bird pollination. Using a target enrichment approach, we obtained 957 loci for 171 accessions to expand the phylogenetic sampling of Neotropical *Costus*. In addition, we performed whole genome resequencing for a subset of 20 closely related species with contrasting pollination syndromes. For each of these 20 genomes, a high-quality assembled transcriptome was used as reference for consensus calling of candidate loci hypothesized to be associated with pollination-related traits of interest. To test for the role these candidate genes may play in evolutionary shifts in pollinators, signatures of selection were estimated as dN/dS across the identified candidate loci. We obtained a well-resolved phylogeny for Neotropical *Costus* despite conflict among gene trees that provide evidence of incomplete lineage sorting and/or reticulation. The overall topology and the network of genome-wide single nucleotide polymorphisms (SNPs) indicate that multiple shifts in pollination strategy have occurred across *Costus*, while also suggesting the presence of previously undetected signatures of hybridization between distantly related taxa. Traits related to pollination syndromes are strongly correlated and have been gained and lost in concert several times throughout the evolution of the genus. The presence of

bract appendages is correlated with two traits associated with defenses against herbivory. Although labellum shape is strongly correlated with overall pollination syndrome, we found no significant impact of labellum shape on diversification rates. Evidence suggests an interplay of pollination success with other selective pressures shaping the evolution of the *Costus* inflorescence. Although most of the loci used for phylogenetic inference appear to be under purifying selection, many candidate genes associated with functional traits show evidence of being under positive selection. Together these results indicate an interplay of phylogenetic history with adaptive evolution leading to the diversification of pollination-associated traits in Neotropical *Costus*.

KEYWORDS

correlated evolution, Costaceae, signatures of selection, Nanopore, targeted enrichment, Zingiberales

Introduction

Recognition of the important role that pollinator interactions play in the diversification of flowering plants traces back to observations by Darwin (1892). Selection can act to mold the characteristics of flowers driven by their predominant or most effective pollinators (Stebbins, 1970), while floral antagonists and abiotic factors may also drive selection in parallel or opposing directions (Van der Niet et al., 2014; Caruso et al., 2019; Dellinger, 2020). The combination of traits (e.g., morphology, color, scent, size, and rewards) associated with particular pollinator groups are often referred to as pollination syndromes (Faegri and Pijl, 1979; Fenster et al., 2004; Rosas-Guerrero et al., 2014). Although the validity of the concept of pollination syndromes has been debated and some studies have not been able to predict pollinators using floral traits either alone or in combination, meta-analyses confirm that there is a higher predictability of pollination syndromes in plants that are distributed in the tropics and associated with bats, bees, or hummingbirds (Rosas-Guerrero et al., 2014; Ashworth et al., 2015). The possibility of both specialized pollination and of frequent shifts in pollination syndromes has been associated with increased opportunities for reproductive isolation (Serrano-Serrano et al., 2017; Fragoso-Martínez et al., 2018) and higher diversification rates (Lagomarsino et al., 2016; Xue et al., 2020). However, in some lineages there is not a clear association between phenotypic traits important for pollination or reproductive mode and rates of diversification (Landis et al., 2018; Moein et al., 2021).

A classic example of plants bearing traits associated with discrete pollinator guilds is the Neotropical section of the genus *Costus* L., a genus of tropical monocots comprising approximately 80 described species (Maas, 1972; Maas-van de Kamer et al., 2016). The genus is currently divided into

a paraphyletic African grade and a large Neotropical clade, the latter of which radiated rapidly after a single long-distance dispersal event from Africa (Specht, 2006; André et al., 2016). Two main pollination syndromes have recursively evolved within the Neotropical *Costus* clade: “mellitophilous” (bee pollinated) species exhibiting flowers with a wide floral tube bordered by a broad labellum that is either yellow or white, typically marked by lateral coloration patterns and a central landing platform bearing a nectar guide or honey mark; and “ornithophilous” (bird pollinated) species, with narrow tubular flowers that are typically yellow, orange, pink, or red in color and lack any lateral markings (Maas, 1972; Specht et al., 2001). Mellitophilous flowers are most often subtended by green, often leafy, bracts while ornithophilous flowers are subtended by bright red, orange, or yellow bracts. Previous phylogenetic studies have estimated multiple shifts in pollination syndromes during the evolution of Neotropical *Costus*, including shifts to mellitophilous pollination syndromes and subsequent regains of ornithophilous flowers (Specht, 2006; Salzman et al., 2015; Valderrama et al., 2020; Vargas et al., 2020), mainly through sympatric speciation events (André et al., 2016). These results suggest that shifts in traits leading to distinct pollination syndromes, together with reproductive isolation *via* reinforcement, may account for the rapid radiation observed in this Neotropical clade.

In addition to ornithophilous species bearing traits that suggest adaptation to hummingbird pollination, there is evidence that some of the ornithophilous traits may also function in a process known as “bee avoidance” (Lunau et al., 2011). Bees are abundant visitors of flowers especially in tropical regions (Bawa, 1990; Renner and Feil, 1993), however they are often found to be less effective pollinators than hummingbirds (Castellanos et al., 2004; Freitas et al., 2006). The bee avoidance hypothesis suggests that certain traits that

we associate with hummingbird-pollinated plants may have evolved to discourage visitation by nectar-robbing insects rather than to attract hummingbirds (Raven, 1972; Lunau and Maier, 1995). Rodríguez-Gironés and Santamaría (2004) showed that bees have limited ability to distinguish red-reflecting colors from the background, resulting in their inability to see red or darkly pigmented orange flowers. This results in their exclusion from red-reflecting flowers, creating a private channel of communication with birds and leaving avian pollinators free of competition for the nectar food source. The potential for bee avoidance behavior has been documented for *Costus*; in polymorphic populations of *Costus arabicus* L. bearing white and dark pink flowers, Bergamo et al. (2016) demonstrated that bees preferentially visited white flowers while hummingbirds show no preference for white or pink flowers. Given the number of evolutionary events in which bird pollination has evolved from bee pollination, Neotropical *Costus* provides an interesting case for further investigation into the evolutionary drivers of trait evolution, whether the traits actively attract effective “specialist” pollinators or prevent visitation by less effective generalists (Schemske and Horvitz, 1984; Thomson, 2003).

Another dimension of selective pressure on the floral morphology of Neotropical *Costus* comes in the form of anti-herbivory strategies. Many *Costus* species have evolved the ability to attract ants *via* the production of nectar by extrafloral glands located on the inflorescence bracts. The presence of ants is a well-documented strategy against herbivory in many plant species (e.g., Pringle et al., 2012), and has been shown to increase the viability of seeds and thereby the fitness of tropical plants by deterring seed herbivory (Rosumek et al., 2009). An example of this is seen in *Costus woodsonii* Maas, where ants deter flies (*Euxesta* sp.; Ulidiidae) from laying eggs on or within the bracts, thereby reducing larval predators that feed on the fruits and damage tissues which are critical for dispersal (arils) and reproduction (seeds) (Schemske, 1980, 1982). The presence of ants could also impact pollination, as bees tend to avoid flowers that are covered by ants in other Neotropical lineages (Nogueira et al., 2021). Similarly, flies are red-sensitive insects that have receptors with spectral sensitivity in those wavelengths (Lunau, 2014; Coimbra et al., 2020), suggesting the possibility that dark pigmentation traits associated with ornithophilous flowers could also impact interactions with egg-laying flies and larval predation. The evolution of *Costus* inflorescence morphology could thus reflect an interplay between selective pressures related to pollination biology and herbivory, influencing trait evolution and the tempo and mode of diversification across this tropical lineage.

Given the diversity of floral and inflorescence morphologies associated with key ecological and environmental interactions, *Costus* provides an outstanding model to study the genetic and genomic mechanisms underlying convergent evolution of organismal interactions, including pollination syndromes. Evolutionary developmental studies have suggested that floral

developmental genes such as *AGAMOUS* (Almeida et al., 2015), *PISTILLATA* (Bartlett and Specht, 2010), and *YABBY2/5* (Almeida et al., 2014; Morioka et al., 2015), as well as the gene regulatory networks in which they participate, may be involved in floral morphological and developmental shifts underlying the convergent evolution of pollination syndrome morphologies (Mendoza et al., 1999; Barrio et al., 2010). These genetic mechanisms may provide keys to understanding the processes that underlie the rapid radiation and diversification inherent to Neotropical *Costus*. Previous genome-level phylogenetic analyses of Neotropical *Costus* (Valderrama et al., 2020; Vargas et al., 2020) demonstrate challenges in recovering species-level resolution due to rapid radiation of morphological features defining species without concomitant divergence of molecular markers. By investigating the evolution and diversification of whole genomes and of candidate genes known to be involved in the evolution of morphological traits that are associated with species-level diversification of *Costus*, we may be able to detect selective pressures that drive the rapid radiation of this lineage and determine if similar mechanisms are responsible for the various transitions among observed morphological innovations. Examples in which trait diversity has been correlated with genomic signatures of selection to investigate mechanisms underlying trait evolution include local adaptation to coastal environments in the Australian Groundsel (*Senecio lautus* Willd.; Asteraceae; Roda et al., 2013) and adaptation to alpine environments in snapdragons (*Antirrhinum* L.; Plantaginaceae; Durán-Castillo et al., 2022).

Beyond looking at selection across individual genes, recent studies have shown evidence of genomic-level changes being driven explicitly by associations with pollinators; examples include higher genetic differentiation (F_{ST}) in bee vs. vertebrate pollinated populations with increased geographic distance (Dellinger et al., 2022), and observed allele frequency changes following generations of bumble bee pollinated treatments when compared with the control (Frachon and Schiestl, 2021). Additionally, linkage disequilibrium (LD) has been observed to decay more rapidly with higher rates of outcrossing due to increased effective recombination (Morrell et al., 2005; Zhu et al., 2016), with self-incompatibility (i.e., forced outcrossing) leading to even more rapid decay as observed in the tea plant (*Camellia sinensis* [L.] Kuntze; Theaceae; Niu et al., 2019). Bird pollination may increase reproductive success as compared to bee pollination (Abrahamczyk et al., 2022), with bird pollination possibly leading to an increase in the rate of outcrossing (Whelan et al., 2009) and thus increased LD decay. Given that positive selection on certain traits can also result in elevated LD, as observed in a floral homeotic mutant of *Aquilegia* L. (Ranunculaceae; Cabin et al., 2022), measures of LD across closely related species can indicate genomic-scale selection in action.

In this study we provide a genome-wide investigation into the genetic underpinnings of morphological traits associated

with shifts in pollinators, as well as conducting a deeper investigation into many candidate genes associated both with specific floral traits and with correlated trait evolution resulting in floral syndrome phenotypes. Specifically, we set out to: (1) provide an updated phylogenetic hypothesis from a comprehensive sampling of the diversity of Neotropical *Costus*, using a target enrichment approach with custom designed genomic baits (Valderrama et al., 2020). We increased the sampling of *Costus* diversity over previous studies by 84 accessions to further test species delimitations, associate biogeographic patterns with morphologic diversity including polymorphism within species, and estimate relationships for newly described taxa. We inferred a robust phylogeny for 66 species (represented by 70 OTUs) using concatenation and coalescent-based species tree methods. (2) Investigate morphological changes associated with pollination syndromes across the genus, to test if inflorescence traits associated with pollination and/or defense against herbivory are constrained by the phylogenetic relationships of Neotropical *Costus* species, and if those traits evolved in a correlated manner. Furthermore, we test if different diversification rates are attributable to traits related to pollination syndromes. (3) Identify signatures of selection across putative neutral loci and candidate genes underlying morphological changes. Genes of interest were identified in the literature as influential for floral development and associated with traits that together generate classic pollination syndromes to test for structural variation associated with “mellitophilous” or “ornithophilous” flowers across the diversity of Neotropical *Costus*. (4) Compare genome-wide signatures across different pollinator types. We obtained Illumina short reads from whole genome sequencing for a subset of 20 closely related species with contrasting (bee vs. hummingbird) pollination syndromes. These data were used to identify gene regions that could be associated with corresponding shifts of floral morphology that were not included in the target bait set; the genome-wide data were also used to investigate relationships among targeted taxa outside of a purely bifurcating framework using a phylogenetic network given that some were of known hybrid origin.

Materials and methods

The phylogenomic approach presented here is based on previous work using a targeted enrichment strategy designed to capture nuclear loci specifically for *Costus*, resulting in a robust topology of the Neotropical *Costus* radiation (Valderrama et al., 2020). Targeted regions were selected using genome assemblies for *Costus spicatus* (Jacq.) Sw. and *Costus longibracteolatus* Maas submitted to the Phyluce pipeline (Faircloth et al., 2012; Faircloth, 2016) with customizations to avoid repetitive and potentially non-homologous regions, removal of plastid genomes and maintaining GC content of the regions between

37 and 55% (details in Valderrama et al., 2020). The baits included target loci identified in previous studies as being single-copy and useful for phylogenomic inferences within the order Zingiberales (240 regions from Sass et al., 2016; 47 regions from Carlsen et al., 2018) as well as 235 loci for genes known to be involved in the evolution of floral development, yielding a total bait set of 1,521 loci and over 1 million base pairs. Custom probes were synthesized in 100 mers with a 20 K design by myBaits (Arbor Biosciences, Ann Arbor, MI, United States) with 3× tiling. Sampling was expanded from Valderrama et al. (2020) to include an additional 31 species (distributed in 76 accessions) of the Neotropical radiation of *Costus* for a total of 61 species (171 accessions) of the c.77 described species (Supplementary Table 1). These accessions were selected to represent morphologic and biogeographic variation within described species. Additionally, 48 accessions representing the major African *Costus* lineages and Neotropical Costaceae were included as outgroups.

Extraction and library prep

DNA was extracted from either silica-dried leaf material collected in the field or from herbarium specimens with protocols that use SDS detergent (Edwards et al., 1991; Konieczny and Ausubel, 1993). Size distributions of the extracted fragments were visualized on a 1% agarose gel and concentrations were quantified with a Qubit 3.0 Fluorometer (Life Technologies, Grand Island, NY, United States). When the majority of fragments for any given sample were longer than 350 bp, a Covaris E220 evolution Focused-ultrasonicator (Covaris, Woburn, MA, United States) was used following manufacturer’s protocols to obtain an average fragment size of 350 bp. Double-sided size selection was performed with size selection beads using a homemade solution of Carboxyl-modified Sera-Mag Magnetic Speed-beads (Thermo Fisher Scientific, Fremont, CA, United States) in a PEG/NaCl buffer (Rowan et al., 2017). The size selection protocol was simplified to only discard smaller fragments for samples from herbarium specimens that yielded low DNA concentrations largely comprising short fragments.

Dual-indexed Libraries were prepared with the KAPA Hyper Prep kit following the manufacturer’s recommendations with 500 ng of DNA fragments around 350 bp in length. Modifications were made to reduce the volume of the reactions to 1/5th of the manufacturer’s protocol shared by Lydia Smith at the Evolutionary Genetics Laboratory at UC Berkeley (protocol¹). Indexed libraries were pooled (6–10 libraries per reaction) and enriched following the manufacturer’s instructions (myBaits Manual v4.01, Arbor Biosciences, Ann Arbor, MI, United States) with a hybridization temperature of

¹ <https://osf.io/fkj2x>

65°C for 24 h. To improve the performance of the blocking oligos, the mix supplied with the baits (i.e., Roche Universal Blocking Oligo Kit and SeqCap EZ Developer Reagent) was substituted with plant C0t-1 DNA (Portik et al., 2016). Pooled libraries were enriched, amplified and sequenced on a lane of NovaSeq SP 150 PE in the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley after final quantification and pooling.

Assembly and alignment

Low quality bases as well as adapters were removed from the reads using TrimGalore² (Martin, 2011) and normalized to 100× coverage using bbnorm within the BBtools package.³ Reads were mapped and assembled to the 1,521 target loci using HybPiper v1.3.1 (Johnson et al., 2016) with default settings. Paralog sequences for the assembled loci were retrieved with HybPiper. The resulting contigs were combined with data from Valderrama et al. (2020) and loci with paralog warnings obtained for more than 5% of the accessions in recovered loci were excluded from downstream analyses. Alignments were estimated with MAFFT v7.271 (Katoh and Toh, 2010) with the iterative refinement method incorporating local pairwise alignment information and with a gap opening penalty of 10 (maximum iterations set to 10,000) and poorly aligned bases and spurious sequences removed with Trimal v1.4.1 (Capella-Gutiérrez et al., 2009) with the automated1 option. Only exonic sequences were used because including introns resulted in poor alignments.

Phylogenomic inference

Gene trees were estimated using IQTree v2.1.2 (Minh et al., 2020), using the ModelFinder (Kalyanamoorthy et al., 2017) option (-m MFP) to find the best DNA substitution model followed by maximum likelihood inference to generate a tree for each partition. TreeShrink v1.3.7 (Mai and Mirarab, 2018) with default values for the species mode ($a = 0.05$, $b = 5\%$) was used to detect irregularly long branches in the gene trees. The TreeShrink algorithm identifies significantly long branches and removes them in the respective trees and alignments using the distribution of branch lengths for each individual within the gene trees. All loci purposed for phylogenetic inference (excluding the floral development genes) were concatenated and a GTR + G substitution model was fit to each gene. Before finding the best tree, partitions corresponding to each gene were merged using the greedy heuristic algorithm (Lanfear et al., 2014). To prevent the exploration of partition merging schemes

from becoming computationally intractable, the relaxed cluster algorithm (rcluster option; Lanfear et al., 2014) was set to examine only the top 10% of the schemes. The ultrafast bootstrap approximation (Hoang et al., 2018) combined with the single branch SH-like approximate likelihood ratio test (Guindon et al., 2010), each with 10,000 replicates, was used to assess support for the resulting branches.

To account for discordance among gene trees, the species tree was estimated with ASTRAL v5.6.3 (Zhang et al., 2018b) using gene trees obtained from IQtree for each locus. The R packages treeio v1.10.0 (Wang et al., 2020) and ggtree v2.0.4 (Yu et al., 2017) were used to plot the quartet support values estimated with ASTRAL on the resulting topology using the -t2 output option. Phytools v0.7-90 (Revell, 2012) function cophylo was used to compare the concatenation and coalescent-based species trees.

Trait scoring and frequency

Features of the inflorescences identified to be related to pollination syndromes were scored as binary traits for all included taxa, using published literature for recognized species and including data from proposed species to be included in an updated monograph of the new world Costaceae (Maas, 1972, 1977; Maas et al., submitted⁴; **Supplementary Table 2**). The traits corresponding to pollination syndrome are: [1] **labellum shape** (tubular/spreading), which directly corresponds to observations of flower visitors, with hummingbirds visiting flowers with a tubular labellum and bees visiting flowers with an horizontally spreading labellum (Kay et al., 2005); [2] **bract color** (red/green), including those with bright orange or yellow bracts in the red category, e.g., *Costus lasius* Loes. and *Costus wilsonii* Maas; [3] **bract appendages** (present/absent), with four species (*Costus barbatus* Suess., *Costus callosus* Maas & H.Maas, *Costus curvibracteatus* Maas, and *Costus juruanus* K.Schum) recognized as having extensions of the bract apex that resemble appendages but are not exactly the same anatomical structure described by Maas (1972). Because extensions can play a similar role in the appearance of the inflorescence, species with extensions were coded as present for appendages; [4] **“honey mark”** or nectar guide (present/absent), as a colored mark located in the center of the labellum; and [5] **labellum stripes** (present/absent), with those coded as present having lateral markings across the labellum that contrast with the ground color. In addition, we also coded a trait that indicated the overall appearance of the inflorescence, [6] **functional color** (red/non-red). In many cases, the color of the bracts is

² <https://github.com/FelixKrueger/TrimGalore>

³ sourceforge.net/projects/bbmap/

⁴ Maas, P. J. M., Maas-van de Kamer, H., Andre, T., Skinner, D., Valderrama, E., and Specht, C. D. (Submitted). A revision of the neotropical Costaceae Nakai: a sixty-year struggle. Part 1: new species. *Phytokeys*.

completely hidden by the presence of appendages (e.g., *Costus leucanthus* Maas or *Costus sinningii* Rusby) such that the inflorescence presents as the color of the appendages. This is likely to influence interactions with pollinators, for example creating a “functionally green” inflorescence even though bracts are scored as red. In this case, the functional color is coded as non-red to reflect the color of the appendages.

Additional traits potentially associated with defense against herbivory, but also relevant for pollination, are: [7] **callus** (present/absent), referring to the extrafloral nectar producing gland located on the abaxial side of inflorescence bracts; and [8] **inflorescence indument** (present/absent), referring to the presence or absence of indument on bracts and/or bract appendages. All polymorphic characters were coded as ambiguous (both states) and our scoring represents the known variation within each species.

The UpsetR v1.4.0 package (Lex et al., 2014) was used to visualize the subset of observed combinations of characters found in our sampling and the frequency of those combinations. Traits were variable for some taxa and the weight was divided into the possible combinations within each lineage (maximum of four combinations due to two variable traits). To prevent including biologically unrealistic combinations, all combinations of characters that were only observed because two traits were polymorphic for the same lineage were removed. Also, four lineages with three ambiguous/polymorphic characters each (*C. arabicus* L., *Costus guanaiensis* Rusby, *Costus lima* K.Schum., and *C. longibracteolatus*) were removed from the analysis.

Comparative methods

Individuals of the same species that formed monophyletic groups were pruned to leave a single accession per species. *Costus acreanus* (Loes.) Maas, *C. comosus* (Jacq.) Roscoe, and *C. leucanthus* accessions were non-monophyletic (see Figure 1 and section “Results”), therefore a single accession was kept per clade (or lineage) to encompass the evolutionary history within each named taxon. Before applying comparative methods, the phylogenetic signal of each analyzed trait was estimated with the D statistic of phylogenetic dispersion (Fritz and Purvis, 2010). Both null hypotheses of the characters evolving following a random (no phylogenetic signal) or a Brownian model (strong phylogenetic signal) were assessed for each binary character (100,000 permutations) using the phylo.d function of the R package caper v1.0.1 (Orme et al., 2013). To reconstruct ancestral character states, stochastic character mapping (Huelsenbeck et al., 2003) was used with models including equal and different transition rates for each binary trait using the make.simmap function of phytools v0.7-90 (Revell, 2012). One thousand stochastic character maps were summarized to estimate posterior probabilities of the ancestral

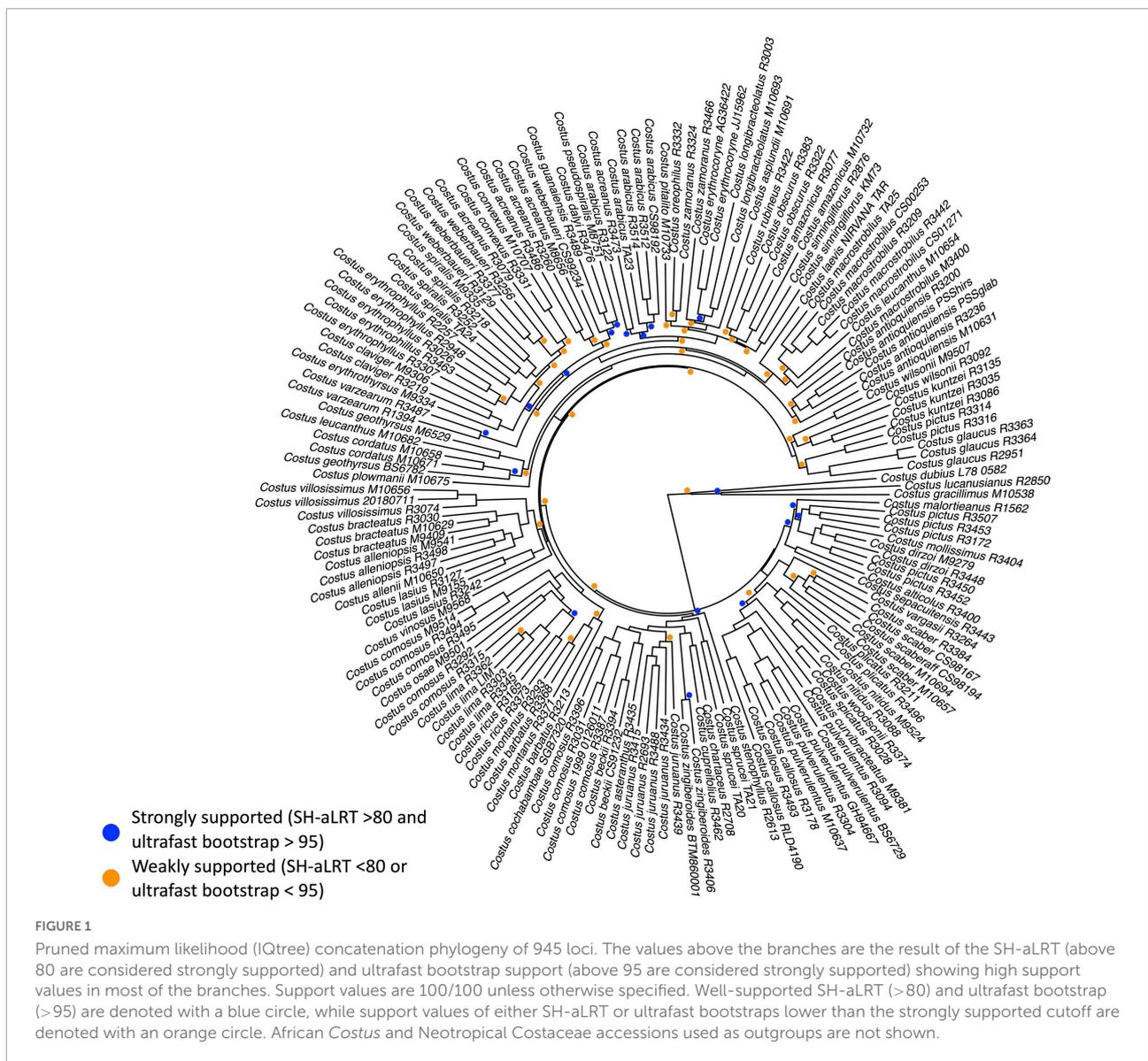
states of each character. The equal and different rate models were compared with a likelihood-ratio test. Because most of the comparative methods require ultrametric trees, the phylogenetic tree from the concatenated analysis was transformed using penalized likelihood with correlated rate variation among branches (Kim and Sanderson, 2008) using the chronos function in the R package ape v5.5 (Paradis and Schliep, 2019).

The Pagel (1994) analysis of correlated evolution was used to test if models of independent evolution vs. dependent evolution were more appropriate for each possible pair of traits given our data; analyses were run with the fitPagel function (fitMk method) available in the R package phytools. Models with equal and different transition rates were tested for each comparison. A Bonferroni (1935) correction was applied to the *p*-value corresponding to the Chi-squared test of the likelihood ratio test to account for the inclusion of each trait in seven comparisons.

To explore if diversification rates are different between lineages with different pollination syndromes, models with constant diversification rates vs. Binary State Speciation and Extinction (BiSSE; Maddison et al., 2007) models were compared with different diversification parameters set for lineages with tubular vs. spreading labellum (a proxy for ornithophilous and melittophilous pollination syndromes). These results were compared to Hidden State Speciation and Extinction models (HiSSE; Beaulieu and O’Meara, 2016) that account for potential distinct diversification regimes unrelated to pollination syndromes by fitting three distinct models, one with a hidden trait accompanying pollination syndromes and two pollination-independent models. For the two trait-independent models, the first included two hidden states (CID-2 HiSSE model) and the second included four hidden states (CID-4 HiSSE model). All models were fitted with the hisse v2.1.6 R package (Beaulieu and O’Meara, 2016) keeping extinction parameters linked.

Transcriptome analyses

Fresh material (whole flowers, inflorescences, and leaves for *C. spicatus*) was flash frozen using liquid nitrogen followed by storage at -80°C . Total RNA was extracted following the protocol of Yockteng et al. (2013), cleaned with homemade ampure beads (Rowan et al., 2017), and quantified with a Nanodrop and Qubit 3.0 Fluorometer (Life Technologies, Grand Island, NY, United States). For short-read Illumina sequencing, RNA was sent to the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley (Berkeley, CA, United States) for their full service library preparation method using ribodepletion and sequencing on an Illumina NovaSeq 2 \times 150. Raw Illumina reads were adaptor-trimmed and filtered for low quality bases using fastp v0.20.1 (Chen et al., 2018). Long reads for leaf and flower RNA were generated for cDNA libraries constructed using the cDNA-PCR Barcoding library kit (SQK-PCS109 and



SQK-PBK004) with 100 ng of total RNA for each sample and sequenced on a single Oxford Nanopore minION flow cell. The resulting fast5 files were used to call fastq files using guppy basecaller v3.4.1 (Oxford Nanopore) with a minimum quality score of seven. Fastq files were then demultiplexed and adaptor-trimmed using porechop v0.2.4.⁵ Nanopore reads were error corrected prior to transcriptome assembly using FMLRC v1.0.0 (Wang et al., 2018) after first building a FM index using the cffq function in MSBWT v0.3.0 (Holt and McMillan, 2014). The error-corrected reads were then filtered using seqkit v0.12.0 (Shen et al., 2016) to discard reads shorter than 800 bp.

A *de novo* transcriptome assembly was generated to serve as a reference to identify candidate genes of interest for

downstream analyses using Trinity v2.11.0 (Grabherr et al., 2011; Haas et al., 2013) using both the cleaned Illumina short reads and Nanopore long reads by implementing the long-reads flag. This approach has been used previously in several studies to generate full length or nearly full length gene contigs in the reference transcriptome assembly to map Illumina short reads to for differential gene expression analysis (Landis et al., 2020; Zhang et al., 2022). Contigs smaller than 250 bp in length were removed with seqkit v0.12.0 (Shen et al., 2016). The assembled transcriptome was annotated using Trinotate v3.2.2 (Bryant et al., 2017) and Transdecoder v5.5.0 (Haas et al., 2013). Contigs were compared to the SwissProt database with ncbi-blast v2.5.0 (Camacho et al., 2009) using BLASTX for nucleotide sequences of contigs and BLASTP for translated amino acids of each contig. Completeness of the assembled transcriptome was

⁵ <https://github.com/rwick/Porechop>

determined by using the embryophyta Benchmarking Universal Single-Copy Orthologs (BUSCO) v3.0.2 (Simão et al., 2015) and summarized in terms of number and size of contigs with NanoPlot (De Coster et al., 2018).

Genome sequencing and analysis

A reference draft genome of *Costus bracteatus* Rowlee was assembled using both Illumina short reads and Nanopore long reads to be used as reference for SNP calling for downstream analyses. Fresh leaf material of *C. bracteatus* R3030 was collected and immediately shipped on ice overnight to California State University East Bay. Leaves were stored at -80° upon arrival until further use. DNA for Illumina sequencing was extracted using the GeneJET Plant Genomic DNA Purification Mini Kit (Thermo Fisher Scientific) following the manufacturer's instructions as described in Grunder (2021) and quantified using a NanoDrop (Thermo Fisher Scientific). DNA was sent to the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley (Berkeley, CA, United States) for library preparation following in-house DNASeq protocol for the Illumina NovaSeq 2×150 to generate approximately 80 GB of sequencing data. Raw Illumina sequencing reads were adaptor-trimmed and filtered for low quality bases using fastp v0.20.1 (Chen et al., 2018). Due to the read depth generated and the unevenness of sequencing across the genome, the cleaned reads were normalized to $100\times$ coverage using the bbnorm function in BBtools (see text footnote 3). Illumina genome sequencing was generated for an additional 19 accessions of *Costus* (Supplementary Table 3) following the same extraction and sequencing protocols. A total of 80 GB of data was targeted for each accession and processed using the same methods described above. The 20 species used for Illumina sequencing were initially identified as sister species with different pollination syndromes. Unfortunately, the most recent phylogeny of Valderrama et al. (2020) showed that many of these species were no longer sister species. However, the genomic data generated still allows insights into genetic changes associated with different pollinator syndromes.

DNA for Nanopore sequencing was extracted using a modified SDS protocol with ampure beads (Schalamun et al., 2019). After extraction, the DNA was further cleaned with a 24:1 chloroform:isoamyl alcohol step, followed by ethanol precipitation with ammonium acetate. Cleaned DNA was then passed through the Qiagen DNAeasy PowerClean Pro Cleanup Kit (Qiagen) and resuspended in 50 μ L of sterile water. Concentration and quality of the final DNA was performed on a Nanodrop. Genomic Nanopore libraries were constructed using the Genomic DNA by Ligation kit (SQK-LSK109) using 1–1.5 μ g of clean DNA as template for library preparation. Each completed library was sequenced on a minION flow cell for 72 h. A total of three flowcells were sequenced, generating

39.34 GB of raw sequencing data. Fast5 files were basecalled with a GPU instance of guppy basecaller v6.0.0 + ab7925058 (Oxford Nanopore) with a minimum quality score of five. Fastq files were then adaptor trimmed using porechop v0.2.4 (see text footnote 4). Nanopore reads were error corrected prior to genome assembly using FMLRC v1.0.0 (Wang et al., 2018) after first building a FM index using the cffq function in MSBWT v0.3.0 (Holt and McMillan, 2014). A hybrid genome assembly approach was undertaken with MaSuRCA v3.4.2 (Zimin et al., 2013) to utilize both the Illumina short reads and the error-corrected Nanopore long reads. Default parameters were used in the assembly except the inclusion of $35\times$ coverage of the long reads. The final assembly was polished using POLCA (Zimin and Salzberg, 2020) as implemented in MaSuRCA v3.4.2. Without flow cytometry analyses, the genome size of *C. bracteatus* was estimated using GenomeScope (Vurture et al., 2017) with the Illumina reads as input.

As a way to assess the quality and completeness of the MaSuRCA hybrid genome assembly approach, the adaptor-trimmed Nanopore long reads were used to assemble the genome of *C. bracteatus* with Flye v2.8.3-b1705 (Kolmogorov et al., 2019). The subsequent assembly was polished with the Illumina reads using POLCA as described for the MaSuRCA assembly. The completeness of both genome assemblies was determined by using the embryophyta BUSCO v3.0.2 (Simão et al., 2015) and summarized in terms of number and size of contigs with NanoPlot (De Coster et al., 2018).

Selection analyses

To assess the evolutionary implications of candidate genes associated with phenotypic traits, especially given the lack of quantitative values of gene expression, signatures of selection were analyzed using codeml (Jeffares et al., 2015) in PAML v4.9 (Yang, 2007). Two different data sets were used: a subset of nearly 300 loci from the full capture set and a smaller data set comprising candidate loci generated from the 20 accessions of Illumina genome sequencing. For the larger data set, the longest open reading frame was first extracted for all 1,472 loci that were retrieved from HybPiper and pruned as described above using TransDecoder v5.5.0. The resulting fasta file was then aligned using MAFFT v7.271 (Katoh and Standley, 2013) using the auto and adjustdirectionaccurately flags, followed by trimming with TrimAL v1.4.1 (Capella-Gutiérrez et al., 2009) with the automated1 option. This filtering left nearly 300 loci, which were checked for the appropriate reading frame using AliView v1.28 (Larsson, 2014). Premature stop codons were replaced with NNN using the exportAlignment function in MACSE v1.2 (Ranwez et al., 2011). The final fasta file for each locus was used to generate a gene tree with RAXML v8.2.12 (Stamatakis, 2014) with 100 rapid bootstrap replicates under a GTR + G model of molecular evolution. For each locus with a cleaned alignment

and gene tree, a summary omega value (dN/dS) was calculated with codeml using model = 0.

For the smaller data set, candidate genes associated with flower color, size, symmetry, and fusion were identified from the annotated transcriptome that was assembled for *C. spicatus* (see above). Contigs for known candidate genes associated with flower size included *BIG BROTHER* (Disch et al., 2006), *GASA/GAST1/GEG* (Shi et al., 1992; Herzog et al., 1995; Kotilainen et al., 1999), *LONGIFOLIA* (LNG; Lee et al., 2006), and *AUXIN RESPONSE FACTOR 2* (*ARF2*; Ellis et al., 2005). Genes associated with symmetry include *RADIALIS* (Coen et al., 1995; Boyden et al., 2012) and *DIVARICATA* (Almeida et al., 1997). Genes associated with fusion include *AGL6* (Koo et al., 2010), *CUP SHAPED COTYLEDON* (Aida et al., 1997), *PETAL LOSS* (*PTL*; Lampugnani et al., 2012), *RABBIT EARS* (Krizek et al., 2006), *SUPERMAN* (Sakai et al., 1995), and *YABBY* (Ha et al., 2010). Genes associated with flower color include chalcone synthase (*CHS*), dihydroflavonol 4-reductase (*DFR*), flavonol synthase (*FLS*), and flavonoid 3'-hydroxylase (*F3'H*) (Holton and Cornish, 1995). A fasta file of contigs for each of the above genes was used as a reference file during mapping of the genomic Illumina data for 20 species using bwa-mem2 v2.2.1 (Vasimuddin et al., 2019). The resulting SAM files were converted to BAM and then sorted using Samtools v1.12 (Li et al., 2009). The sorted BAM files were then split using BamTools v2.5.1 (Barnett et al., 2011) so that each candidate gene was in a separate file. Consensus sequences for the coding region of each candidate gene was then generated using the dark-matter⁶ python script make-consensus.py which utilized bcftools v1.9 (Li, 2011) and iVar v1.3.1 (Grubaugh et al., 2019). The resulting fasta files for each locus were combined into a single fasta file, followed by alignment with MAFFT, cleaning with TrimAL, removal of premature stop codons, and gene tree inference as described for the large data set. Due to most gene sequences possessing polymorphic bases represented with ambiguity codes, the PAML settings were changed to remove ambiguous sites. For each locus, a summary omega value (dN/dS) was generated with codeml using model = 0 and an omega value for every branch was generated with model = 1. Omega values less than one indicate purifying selection, omega values equal to one indicate neutral evolution, and omega values above one indicate positive selection.

Single nucleotide polymorphism calling analyses

To investigate genomic changes associated with different pollinators, and to more fully explore the evolutionary history of a subset of taxa, SNPs were called for the 20 species

⁶ <https://github.com/acorg/dark-matter>

with Illumina genome sequencing data. Cleaned reads were mapped to the *C. bracteatus* polished MaSuRCA assembly using bwa-mem2 v2.2.1 (Vasimuddin et al., 2019). Resulting SAM files were converted to BAM and then sorted using Samtools v1.12 (Li et al., 2009). Duplicates were marked using picard v2.21.2 (Broad Institute, 2019). For each species, SNPs were first called using HaplotypeCaller in GATK v4.1.4 (McKenna et al., 2010) using the ERC flag to create a GVCF (Genomic Variant Call Format) for each species. All GVCFs were then combined using the GenomicsDBImport followed by joint genotyping using GenotypeGVCFs. Resulting VCF (Variant Call Format) files were then combined into a single VCF using the GatherVcfs function.

The VCF file was initially filtered in VCFtools v0.1.16 (Danecek et al., 2011) using the following parameters: biallelic sites with a maximum missing 0.9, minimum mean depth of 15, maximum mean depth of 200, minimum quality score of 30, minor allele frequency of 0.05, and minor allele count of 3. LD was calculated using PopLDdecay (Zhang et al., 2018a) on the initial filtered VCF for all samples, ornithophilous species only, and mellitophilous species only. The VCF file was further filtered by LD pruning using plink v1.90b6.21 (Purcell et al., 2007) using a r^2 of 0.2 using a 20 kb window with a 10 kb step size. A phylogenetic network was generated in SplitsTree v4 (Huson and Bryant, 2006) using the final filtered VCF file containing 3,432,721 SNPs after converting to a nexus format with vcf2phyllip (Ortiz, 2019). A phylogenetic network allows for conflicting splits due to gene flow, which we know exists due to our sampling of *Costus vinosus* Maas × *allenii* Maas R3353 and *C. wilsonii* × *villosissimus* Jacq. “mellow yellow” R3091, both being suspected hybrids sampled from plants in cultivation (D. Skinner, personal observation). In addition to the phylogenetic network, a principal components analysis (PCA) was conducted using the R package SNPrelate v3.14 (Zheng et al., 2012) coloring species as bird pollinated, bee pollinated, or known hybrids. The first four principal components (PC) were plotted. Lastly, the level of heterozygosity from the final pruned SNP data set was calculated for the three groups: bee pollinated, bird pollinated, and known hybrids using VCFtools v0.1.16 (Danecek et al., 2011).

Results

Phylogenetic relationships

The capture approach generated 945 loci for a total alignment of 751,609 bp with 355,314 distinct patterns, 125,357 singleton sites and 425,080 constant sites, for 171 accessions representing 61 of a total of ca. 77 species of the Neotropical radiation of *Costus* (see **Supplementary Table 1**). The topology obtained with the concatenation approach is well supported with most branches showing high support values with the

SH-aLRT (>80) and ultrafast bootstrap (>95) (Figure 1 and see Supplementary Figure 1 for the full phylogeny including outgroups). These values are comparatively lower for the lineages comprising the earliest branches of the Neotropical *Costus* radiation. The species trees approach implemented in ASTRAL produced a similar topology but with some differences in the earliest branches (Supplementary Figure 2). The normalized quartet score of the topology obtained with ASTRAL is 46.82%, suggesting high levels of discordance among gene trees. The quartet scores indicate high levels of gene tree conflict in less supported branches and some of the branches with high local posterior probabilities show several gene trees supporting the alternative topologies of each quartet (Supplementary Figure 3). The same well-supported lineages were recovered with both methods and we used the topology inferred from the concatenation approach in subsequent analyses. Most species that were sampled for more than one individual are recovered as monophyletic in the resulting phylogeny, even when considering broad geographic variation or morphological variation. However, some accessions determined as conspecifics comprise independent lineages, such as individuals from western Mesoamerica identified as *Costus pictus* D. Don, that are not related to the accessions of individuals identified as belonging to the same species from eastern Mesoamerica. Our sampling of *C. comosus* and *C. acreanus* reveals at least three independent lineages included in each species, suggesting the presence of species complexes. Similarly, *C. leucanthus* is represented by two lineages highlighting additional hidden diversity within the genus.

Traits combinations and comparative methods

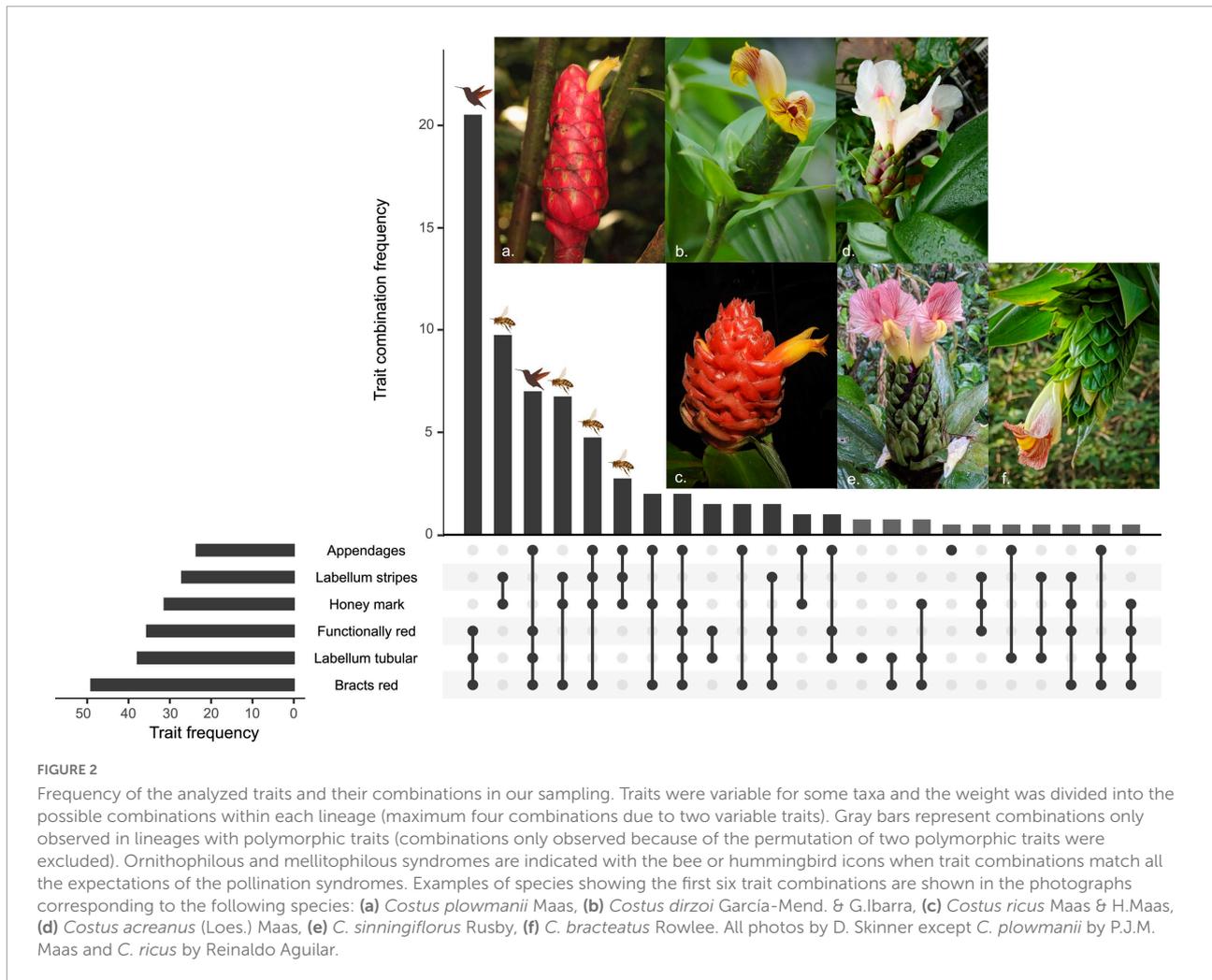
The most prevalent combinations of the analyzed traits match expectations of pollination syndromes (Figure 2). The most frequent combination of traits corresponds to a morphology associated with hummingbird pollination with a tubular labellum (without stripes or honey mark), overall red inflorescences, and no appendages on the bracts (Figure 2a). The second most observed combination agrees with bee-pollinated inflorescences, with a spreading labellum (with stripes and honey mark) and green bracts, and no appendages on the bracts (Figure 2b). The third and sixth most common combinations correspond to hummingbird and bee-pollinated morphologies, respectively, but with appendaged bracts (Figures 2c,f). The fourth most frequent combination of traits matches bee pollination with spreading labellum (with stripes and honey mark) and red bracts, but with functionally non-red inflorescences, the tips of the bracts are green in color (Figure 2d). Red bracts can also be concealed by the green and leafy appendages of the bracts as in the fifth most prevalent combination (Figure 2e), seen in *C. bracteatus*

and *C. sinningiiflorus*, making this morphology functionally identical to that of the sixth most-common combination (compare Figures 2e,f). The other observed combinations do not match all the expectations of pollination syndromes but are less frequent and constitute the tail of a positively skewed distribution.

Comparing the empirical D statistic for each trait with permutations cannot reject the hypothesis of the traits evolving under a Brownian motion model (or with a strong phylogenetic signal) (Supplementary Table 4). The hypothesis that the analyzed traits evolved randomly or with no phylogenetic signal was rejected for all traits ($p < 0.05$) with exceptions for bract color, presence/absence of appendages in the bracts and presence/absence of the callus. The presence or absence of the callus in the bracts is the only trait where the D statistic value is closer to the ones obtained under a model of random evolution (with no phylogenetic structure). The all-rates-different model significantly improved the likelihood scores for how the models fit the transitions of the labellum shape, labellum stripes and appendages in the bracts. For functional color, bract color, presence of the honey mark, callus and inflorescence indument, an equal rates model was favored (Supplementary Table 4).

The stochastic character mapping of the labellum shape indicates that the tubular shape is the most likely ancestral state of Neotropical *Costus* and that the spreading labellum evolved twice and subsequently reverted to tubular shape in several instances (Figure 3; Labellum shape). The other characters related to pollination appear to shift after changes in labellum shape, following the expectations of trait combinations that constitute pollination syndromes. Lineages gaining a spreading labellum subsequently become green in their inflorescence color, acquire labellum stripes, and acquire a honey mark (see Figure 3). The stochastic mapping of the bract appendages does not show the same pattern and suggests that these structures have evolved multiple times in several lineages from ancestors without appendages, independently of other traits and of pollination syndrome. The callus in the bracts has been lost in several lineages independently, while the inflorescence indument has been gained independently in several instances (Figure 3).

The Pagel analyses of correlated evolution show that pollination related traits (labellum shape, functional inflorescence color, honey mark, and labellum stripes) are strongly correlated ($p < 0.00001$) with bract color also correlated ($p < 0.004$) (Figure 4). The labellum shape, thought to be the most important character for pollination syndrome, is only marginally correlated with the evolution of appendages in the bracts. The evolution of the callus and the inflorescence indument, which potentially play a role in defense against herbivory, are not correlated with each other or with the pollination traits, but are both strongly correlated with the presence of appendages on the bracts (Figure 4). The transition rates for the model of correlated evolution of the labellum shape



and the functional color of the inflorescence indicate that the most stable combinations of characters match the pollination syndromes, with the alternative combinations transitioning promptly to a tubular labellum shape and red inflorescence or to a spreading labellum and non-red inflorescence (Figure 5). Note that a non-red functional color can be obtained either by changing to a green bract color or by gaining leafy appendages. In the model of correlated evolution of the labellum shape and the bract appendages, the transition rates of the appendages are higher in the lineages with a spreading labellum, suggesting that this trait is more labile in the lineages with this labellum shape (which is also associated with bee pollination). Similarly, when appendages are present, the transition rates for gaining or losing the callus in the bracts are higher. The lineages with the indument on the inflorescence also show higher transition rates for the presence or absence of the bract appendages (Figure 5).

The model with constant diversification rates has the lowest fit when comparing AICc values. Models with hidden states that are independent of the labellum shape have the best fit, therefore we found no evidence of differential diversification

rates related to the labellum shape and consequently to the pollination syndromes (Supplementary Table 5). The CID4 model had a deltaAIC of 18.21 to the second closest model of the full HiSSE model.

Transcriptome and genome assemblies

After quality filtering and adaptor trimming of the RNA-Seq raw Nanopore reads, a total of 2,291,573 reads with a N50 of 818 bp resulted for the *C. spicatus* whole flower and 2,605,893 reads with a N50 of 822 bp for the leaf material. Illumina sequencing generated 104 million reads for the flower material and 80 million reads for the leaf material (Supplementary Figure 4a). The Trinity transcriptome assembly generated 227,713 contigs with an N50 of 1,795 bp. According to the embryophyta BUSCO library, the *C. spicatus* transcriptome contained 1,335 out of 1,375 complete BUSCOs, with 560 represented as single-copy and 775 represented as duplicated.

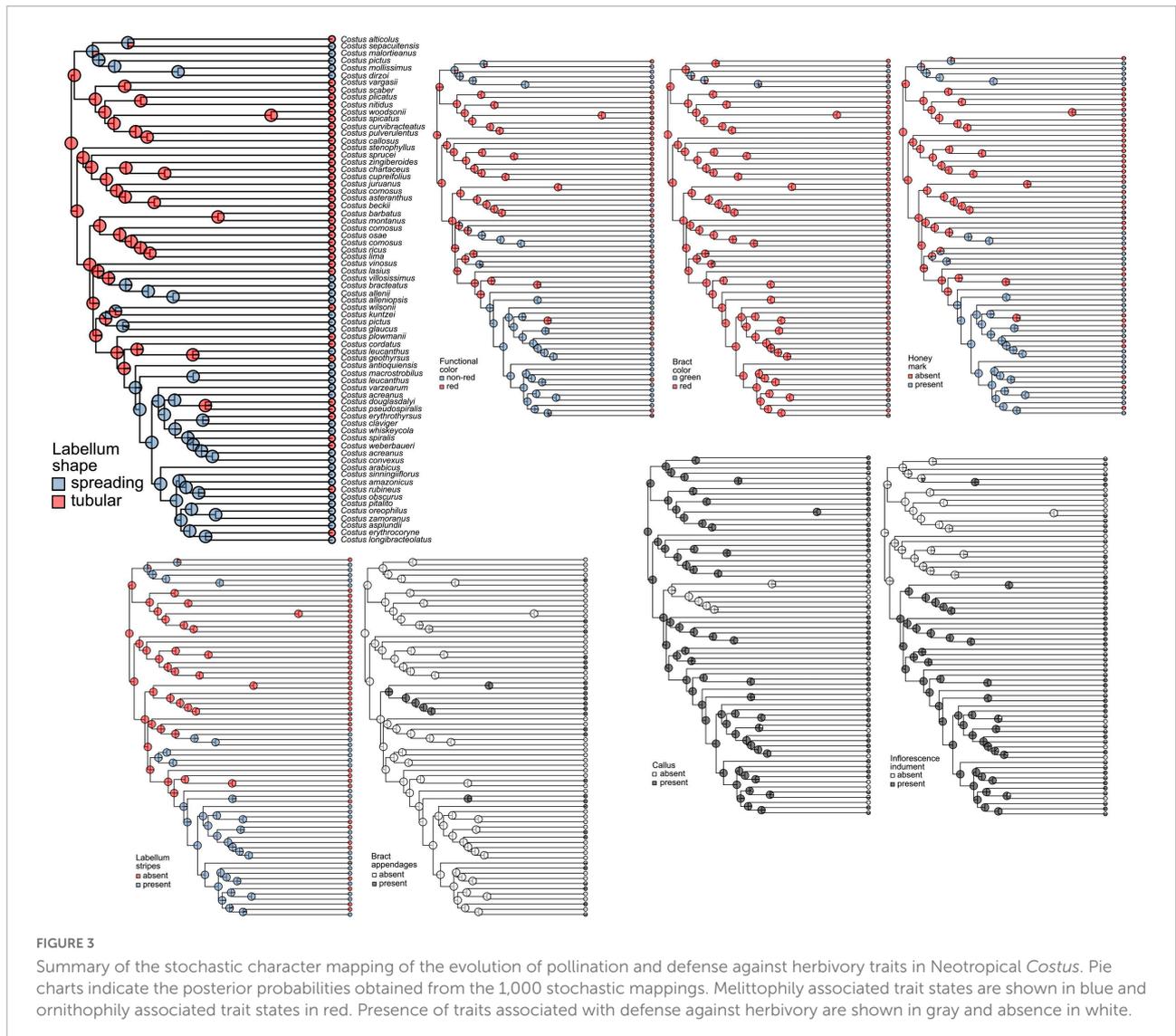


FIGURE 3

Summary of the stochastic character mapping of the evolution of pollination and defense against herbivory traits in Neotropical *Costus*. Pie charts indicate the posterior probabilities obtained from the 1,000 stochastic mappings. Melittophily associated trait states are shown in blue and ornithophily associated trait states in red. Presence of traits associated with defense against herbivory are shown in gray and absence in white.

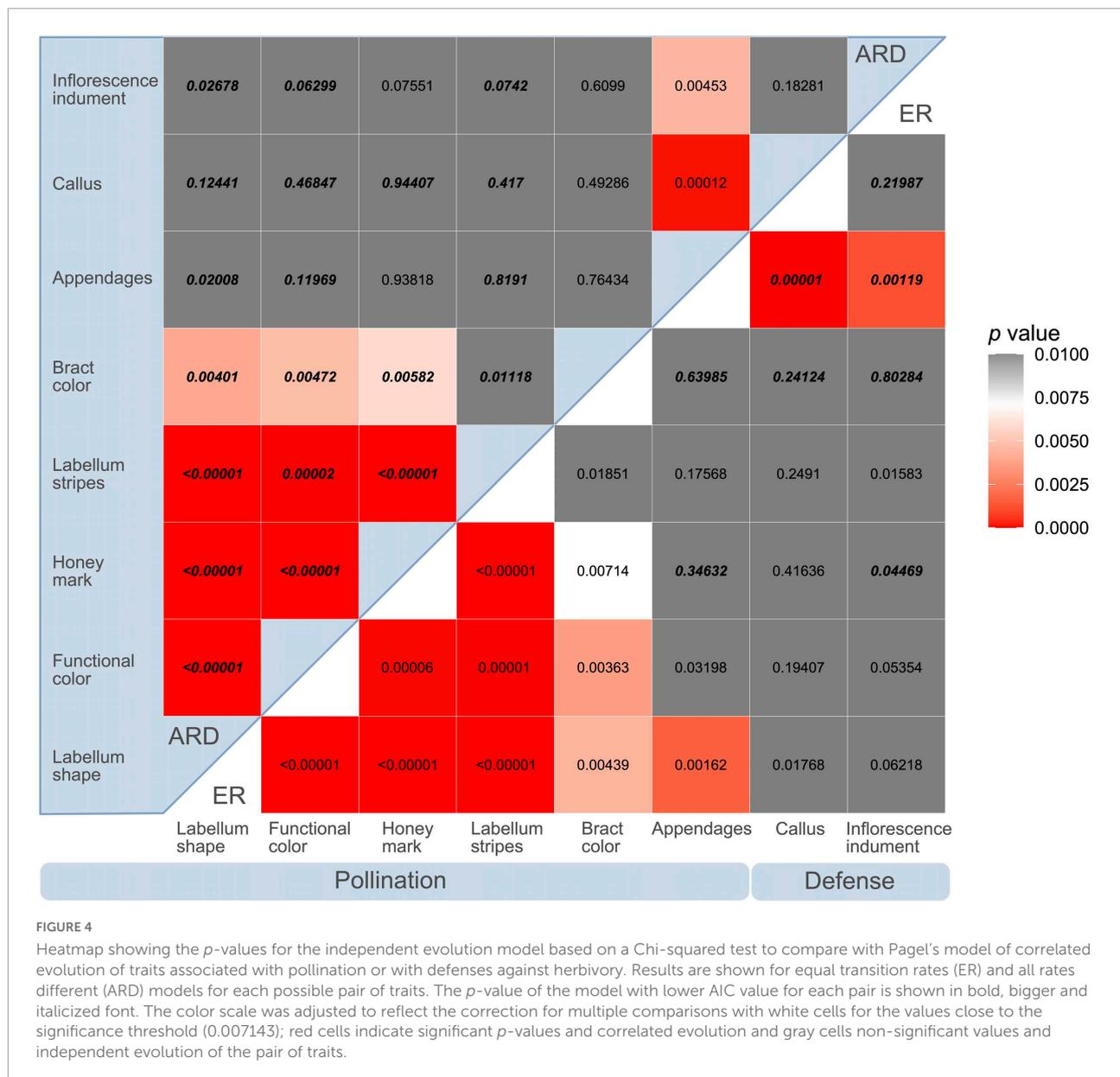
The transcriptome assembly had 18 fragmented BUSCOs and 22 missing BUSCOs.

After filtering, adaptor trimming, and normalization of the genome Illumina data for *C. bracteatus* a total of 133,777,645 reads representing approximately 40 GB of data remained. After filtering the Nanopore data, a total of 13,164,642 reads with an N50 of 5,923 bp comprising 37 GB of data remained. The polished MaSuRCA genome assembly contained 13,709 contigs with an N50 of 142,420 bp with a total assembly size of 958 MB (Supplementary Figure 4b). In comparison, the polished Flye assembly contained 44,677 contigs with an N50 of 156,862 bp with a total assembly size of 617 MB. Based on the Illumina short reads, the estimated genome size of *C. bracteatus* is 834 MB. In terms of completeness, both the polished MaSuRCA and polished Flye assemblies were similar. Prior to polishing, the MaSuRCA assembly contained 1,288 out of 1,375 complete BUSCOs with 1,161 single-copy and

127 duplicated, 33 fragmented, and 54 missing BUSCOs. After polishing the assembly contained 1,291 out of 1,375 complete BUSCOs with 1,163 single-copy and 128 duplicated, with an additional 33 fragmented BUSCOs and 51 missing BUSCOs. In comparison, prior to polishing the Flye assembly contained 1,206 out of 1,375 complete BUSCOs with 1,115 single-copy and 91 duplicated, 97 fragmented, and 72 missing BUSCOs. After polishing the assembly contained 1,297 out of 1,375 complete BUSCOs, with 1,189 single-copy and 108 duplicated, with 30 fragmented, and 48 missing BUSCOs.

Selection analyses

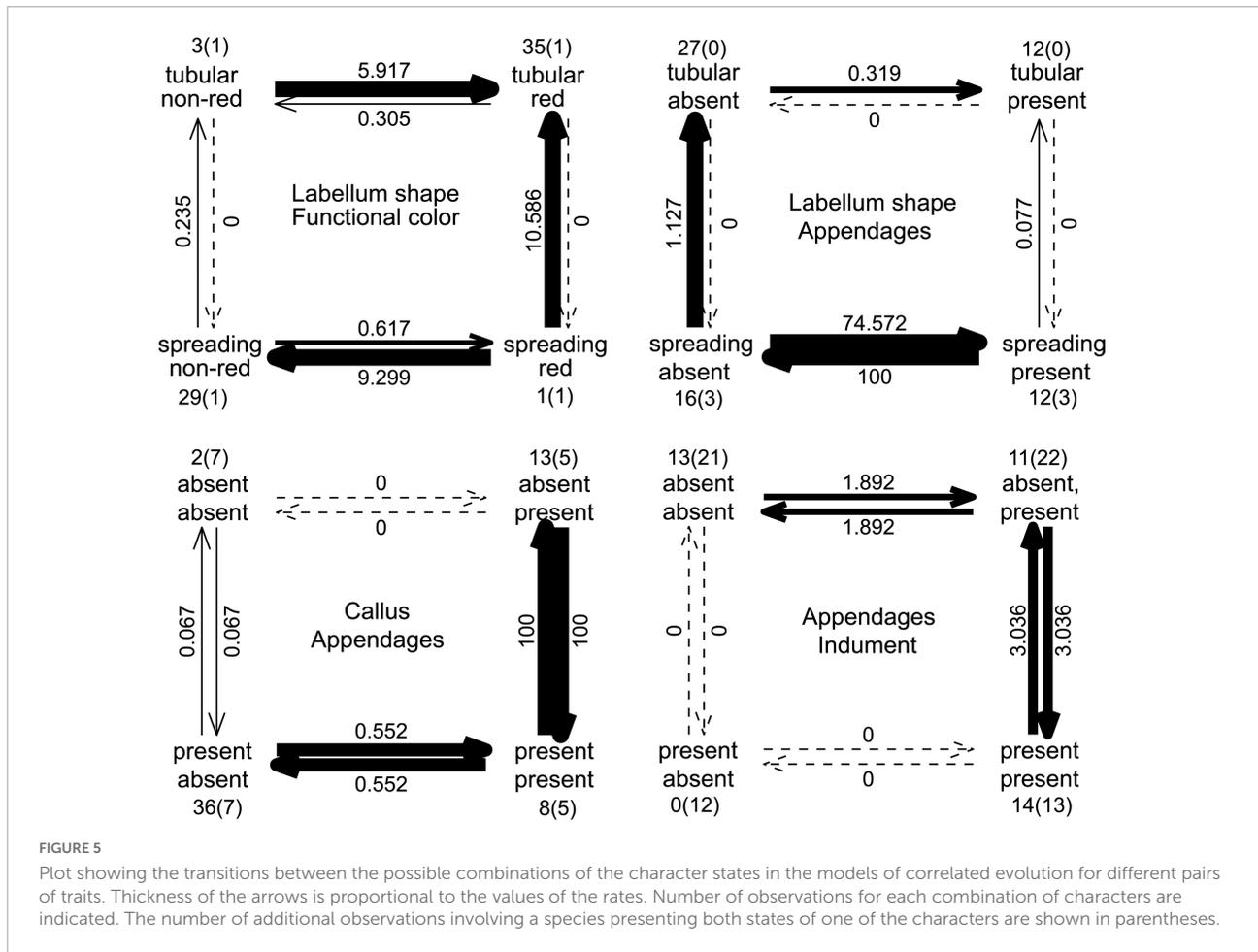
After alignment and filtering of the 957 capture loci, 296 were used for selection analyses (Supplementary Table 6). An additional 44 candidate loci responsible for functional



traits associated with pollination syndromes were used after identification in the *C. spicatus* assembled transcriptome and were extracted from the Illumina sequences using a consensus-calling approach. The breakdown of the capture loci groups were 224 loci developed with Phyluce, 20 loci from Carlsen et al. (2018), 22 loci from Sass et al. (2016), and 32 loci of functional trait genes (Figure 6). The median omega value for the loci used for the phylogeny was 0.503, while the median value for the genes underlying functional traits was slightly elevated with a value of 0.534. Not all of the phylogenetic loci were congruent in terms of selective pressure, with the Sass et al. loci having the smallest median omega value of 0.163 but also the largest omega values recovered. The Carlsen loci showed the most loci exhibiting positive selection with 30 of

the 224 loci (13.4%) having omega values above one. Five of the capture genes associated with functional traits showed evidence of positive selection, which included two loci of *FLS*, two loci of anthocyanin synthase (*ANS*), and one locus of *PTL*.

Out of the additional 44 candidate genes investigated using consensus calling, 11 loci showed evidence of positive selection (omega >1), with five loci having omega values larger than two. The adaptive loci included two loci of *BIG BROTHER*, two loci of *CHS*, one locus of *CUP SHAPED COTYLEDON 2*, one locus of *DIVARICATA*, two loci of *GASA/GAST1*, one locus of *PTL*, one locus of *RADIALIS*-like, and one locus of *SUPERMAN*. Four loci, including one locus of *GASA*, two loci of *LONGIFOLIA*, and one locus of *RADIALIS*-like show a nearly infinite value of omega due to a dS value of zero or

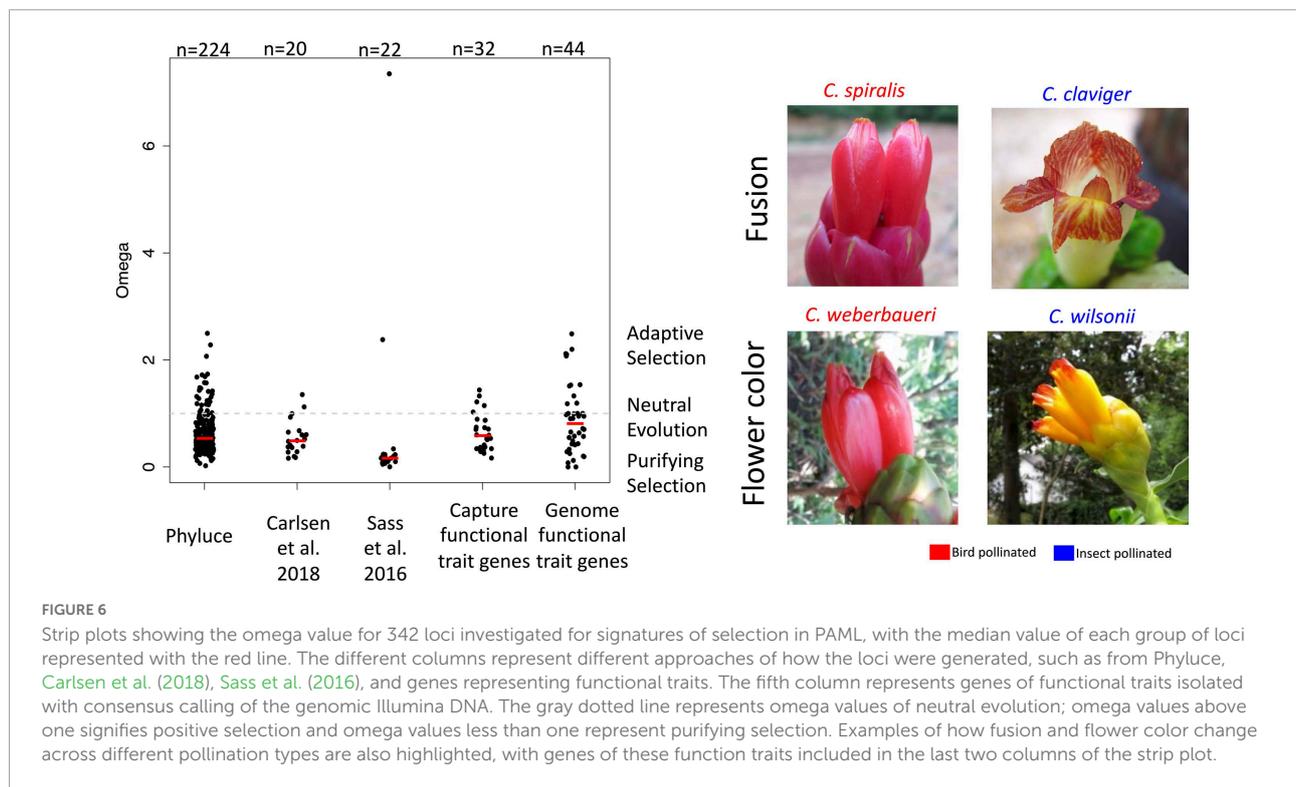


nearly zero (these four loci were not represented in **Figure 6**). Further investigation of these loci by comparing branch-specific omega values and the phenotypic traits scored resulted in a lack of clear patterns. Comparing tubular labellum to the branch specific omega values for *PTL* (**Supplementary Figure 5a**) and *SUPERMAN* (**Supplementary Figure 5b**) showed a slightly elevated omega value for the species with a tubular labellum (median omega = 3.62) vs. those without a tubular labellum (median omega = 3.07). The pattern of selection pressure for *PTL* is more extreme, but far from conclusive. The median omega value of species without a tubular labellum is 0.0001 (with one species, *Costus kuntzei* approaching infinity due to a very low DS value of 0.000005), while five of the species with a tubular labellum have omega values approaching infinity and four species showing evidence of purifying selection (omega < 1) (**Supplementary Figure 5**). The patterns observed between the trait of functionally red and two of the chalcone synthase genes is quite different than for fusion. The median values of omega for *CHS4* (**Supplementary Figure 6a**) are identical between species that are functionally red and those that are not (0.346 vs. 0.346). In contrast, species that are not functionally red show an elevated median omega value (4.887) compared

to those species that are functionally red (3.647) for *CHSY* (**Supplementary Figure 6b**).

Single nucleotide polymorphism analyses

The initial VCF produced from GATK consisted of 115,295,456 raw SNPs. The initial filtering parameters reduced the number of SNPs to 9,992,569. Pruning for LD with an r^2 value of 0.2 left 3,432,721 high quality SNPs. Calculating the LD decay by splitting the 20 samples across bee-pollinated and hummingbird-pollinated species, while excluding known hybrids, showed that the level of LD decreases rapidly and plateaus within approximately 5,000 bp. While both bee and hummingbird-pollinated species show an overall similar pattern, the two groups show differences in the baseline value (**Figure 7**). Bee-pollinated species show a baseline value of LD with an r^2 value of 0.15, while hummingbird-pollinated species show a baseline r^2 value of 0.10. In addition to differences in LD, the two different pollinator groups also display differences in heterozygosity within the variant

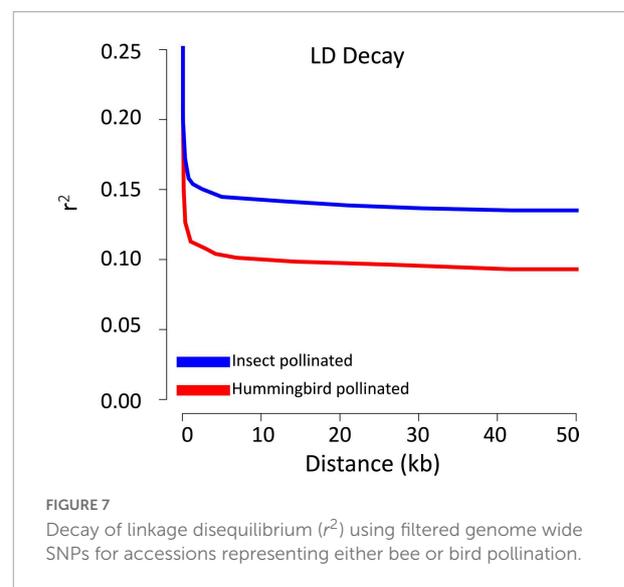


sites. The hummingbird-pollinated species show a median heterozygosity value of 0.131, while the bee-pollinated species show a median heterozygosity value of 0.109. Notably, the two additional species that show conflicting splits in the neighbor network (see below), showed elevated levels of heterozygosity. The hummingbird-pollinated species *C. spicatus* had a heterozygosity value of 0.229, while the bee-pollinated *C. pictus* had a heterozygosity value of 0.217. If these two species are removed from the comparison, then the median heterozygosity value for the hummingbird species is 0.129 compared to 0.106 for the bee-pollinated species. The two known hybrids in the data set have a much higher median heterozygosity value of 0.288, almost double that of all other species.

The neighbor network phylogenetic network shows a similar relationship pattern for the subset of species with full genome data compared to those of the full capture data set (Supplementary Figure 7). However, there are conflicting splits across the genome data that suggest evidence of gene flow *via* hybridization. Two known hybrids in the data set, *C. vinosus* × *alenii* and *C. wilsonii* × *villosissimus* “mellow yellow,” show clear hybridization patterns with *C. villosissimus* appearing to be one of the genetic contributors to these accessions. Other species including *C. spicatus* and *Costus spiralis* (Jacq.) Roscoe which are not of known hybrid origin show conflicting splits between other species. As shown with the larger data set, the hummingbird-pollinated species do not all cluster together and neither do the bee-pollinated species. There is a larger cluster

of exclusively hummingbird-pollinated species including *Costus beckii* Maas & H.Maas, *C. comosus*, *Costus productus* Gleason ex Maas, *Costus stenophyllus* Standl. & L.O. Williams, and *Costus zingiberoides* J.F.Macbr. Other hummingbird-pollinated species including *C. lasius*, *C. spiralis*, *C. spicatus*, and *C. wilsonii* each have bee-pollinated species as their closest relatives.

After reading in the filtered VCF containing 3,432,721 SNPs, a total of 3,104,097 biallelic SNPs were used for the PCA. The top



four PC explained a total of 33.02% of the overall variation with PC1 explaining 9.53%, PC2 explaining 8.91%, PC3 explaining 7.84%, and PC4 explaining 6.74% (Figure 8). Similar to the phylogenetic network, there is no clear grouping of only bee- or hummingbird-pollinated species, with several cases where the closest relatives genetically are from different pollinator groups. The two known hybrids also cluster together, with evidence that one parent is bee pollinated (*C. villosissimus*) and the other is hummingbird pollinated (*C. wilsonii*).

Discussion

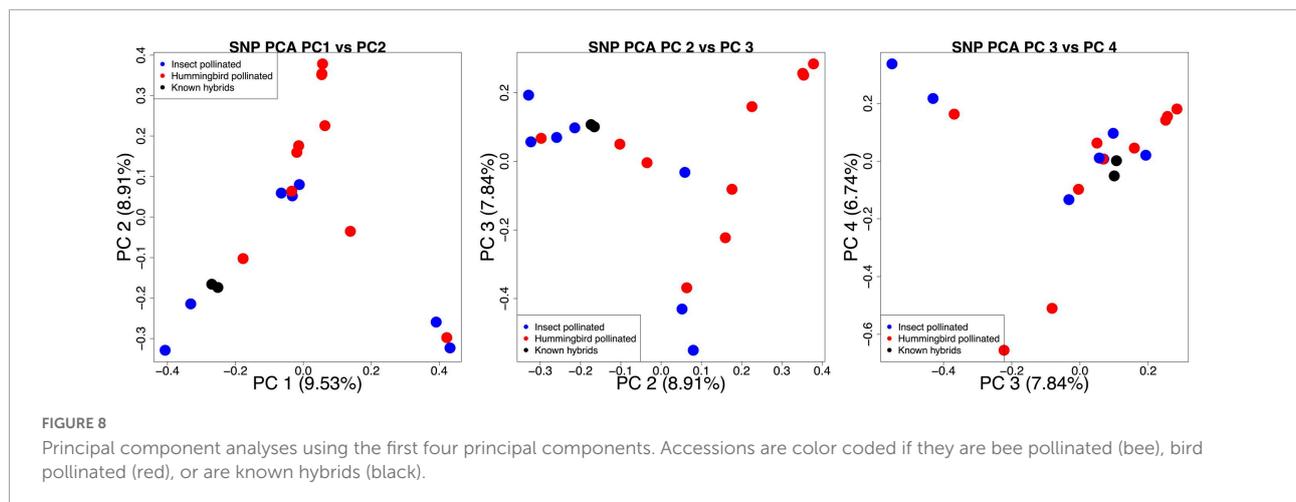
We present here the most complete phylogeny of Neotropical *Costus* including 61 (of c.77) species with multiple exemplars capturing species-level morphologic and biogeographic diversity. Most relationships in the inferred tree are well-supported with the exception of some of the earliest branching events. Most species represented by multiple accessions are monophyletic, though the non-monophyly of some species suggests either cryptic species or the presence of multiple species complexes with ongoing gene flow. In general, trait combinations follow expectations of established ornithophilous and mellitophilous pollination syndromes. One of the more important traits for distinguishing pollination syndromes, labellum shape, appears to have evolved from a tubular to a spreading labellum multiple times followed by corresponding evolutionary shifts in other pollination-associated traits (floral color, nectar guide, lateral markings, and functional inflorescence color). Other traits, such as appendages, appear to be more labile in their evolutionary history. When single traits evolve away from expected pollination syndromes, other traits associated with the suite of traits representing established pollination syndromes appear to rapidly change to conform to a different pollination strategy and converge on the associated pollination syndrome. These rapid transition rates indicate strong adaptation to effective pollination strategies. Indeed, we see signatures of positive selection in several candidate genes associated with floral fusion and pigment production, two traits associated with the development of pollination-associated morphologies. In addition to strong morphological differences between pollination syndromes, we also see genetic differences between individuals of bee and bird pollination such as LD and levels of heterozygosity, which may be associated with outcrossing rates or geographic distance between the two parental genotypes. Even with the presence of strong morphological and genetic signatures of pollination syndromes, we do not recover evidence that the labellum shape impacts rates of diversification of Neotropical *Costus*.

We confirm the idea of a rapid radiation of the Neotropical lineage of *Costus* with short branch lengths obtained for the early nodes of the diversification of the genus in this region (Figure 1; André et al., 2016; Valderrama et al., 2020; Vargas

et al., 2020). The gene tree discordance observed for the normalized quartet scores of the coalescent-based species tree topology can also be attributed to a rapid diversification, leading to incomplete lineage sorting as evidenced in many neotropical plant lineages (Schley et al., 2018; Valderrama et al., 2018; Pérez-Escobar et al., 2021; Torres Jimenez et al., 2021). Alternatively, and not mutually exclusive, the topology and the gene tree discordance could be signatures of hybridization and the resulting introgression over the history of the genus. From the phylogenetic network analysis, we observe evidence of hybridization, both in several known hybrids but also in accessions that previously have not been suspected to be hybrids (Supplementary Figure 7). However, this updated phylogeny allows testing and confirming monophyly when considering geographic and morphological variation within described species. The phylogeny also reveals close relationships of undescribed lineages, clarifying their validity as additional taxonomic units (e.g., *Costus antioquiensis* sp. nov. Maas & H.Maas., *C. oreophilus* sp. nov. Maas & Skinner, or *Costus weberbaueri* Loes.). We further confirm the need for an updated taxonomic revision that accurately reflects the diversity in the genus.

The frequency of combination of traits matching the expectations of two pollination syndromes, bee and bird pollination, as well as the strongly correlated evolution among these characters, points to the validity of the inferred relationship of those pollination syndromes with key traits, like the labellum shape (Maas, 1972; Specht et al., 2012; Maas-van de Kamer et al., 2016). Additional traits such as labellum stripes, honey mark, and inflorescence/bract coloration appear to be acquired by lineages to fine-tune the attraction to pollinators. From the clustering of SNPs genome-wide in the PCA, we do not see clusters of species that exclusively exhibit one pollination syndrome or the other (Figure 8), providing more support that lineages are evolving the necessary traits for effective pollination. The evolution of pollination related traits in *Costus* implies that if the hummingbird morphology evolved to avoid visitation by bees, as posited by the bee avoidance hypothesis (Lunau and Maier, 1995; Lunau et al., 2011), then some lineages subsequently “avoided avoiding bees” by regaining traits that attract bees, including functionally green color. The process of re-evolving traits sometimes involves the layering of structures, such as the acquisition of green appendages that obscure the presence of red bracts in *C. leucanthus* and *C. sinningiiiflorus*. In this case, rather than “losing” red bracts and thereby shifting inflorescence color from red to green, leafy appendages are added creating a functionally green display. Further studies exploring the efficacy of these pigments and cost of their production in the bracts will be required to provide more insights as to the evolutionary history of bract color and the potential selective pressures acting on this inflorescence trait.

We have only limited pollination data for species of *Costus*, with pollinators observed and pollination efficacy measured for



12 species (Kay and Schemske, 2003; Bergamo et al., 2016). Observations of floral visitors of intermediate morphologies described here should be studied in the field to fully understand the selective pressure shaping the morphological variation in this radiation (Bergamo et al., 2016). Monophyletic species like *C. lima* K.Schum. with a distribution ranging from Ecuador to Costa Rica, from sea level to 2,000 m above sea level, possessing morphological variation in the indument and functional color of the inflorescence, constitute outstanding opportunities to test many hypotheses about how traits influence and are influenced by pollination biology in Neotropical plants. Questions that can be addressed include: do floral visitors and herbivory vary over the elevation gradient? How do traits associated with different pollination syndromes and defenses against herbivory actually affect floral visitors and predators along environmental gradients? Objectively measuring the spectral characteristics of the inflorescences and understanding the differences taking into account the visible spectrum of both pollinators and predators could greatly improve our understanding of the communication channels shaping these biotic interactions (Coimbra et al., 2020). Additionally exploring the signals in the UV spectrum (overlooked by our character coding) could also provide valuable information to understand the pollination biology of *Costus* (Bergamo et al., 2016). Studying the morphology of hybrids presenting trait combinations that do not match the pollination syndromes, coupled with observations of floral visitors and quantification of reproductive success, could reveal reinforcement of reproductive isolation with unfit individuals resulting from gene flow between lineages attracting different pollinators (Schley et al., 2022).

The appendages of the bracts are an interesting trait that could have a role in both the attraction of pollinators and defenses against herbivory. Appendages are marginally correlated with labellum shape and hence could play a role in pollination but are not correlated with any of the

other pollination-related traits. The faster transition rates of appendages in species with a spreading labellum suggest that these structures can play a role both in the attraction of bee pollinators but also defending those inflorescences from herbivory, based on the strong correlation with the nectariferous callus that attracts ants that defend the inflorescence from predation. We hypothesize that the presence of certain types of appendages could interfere with the function of the callus or have a function protecting the inflorescences from herbivory. Also, the presence of appendages is correlated with the presence of an indument in the inflorescence, perhaps jointly playing a role against herbivory; the gain or loss of appendages mainly involves lineages with an indument already present on their inflorescences. The combination of traits displayed in *Costus* inflorescences seems to be shaped by the interactions with pollinators and predators forming a complex ecological network (Irwin et al., 2004; Dellinger, 2020). Furthermore, these traits could also play a role in attracting seed dispersers, adding another set of organismal interactions that could potentially modulate selective pressures over inflorescence traits. The complexity and interaction of many traits involved in pollination and defenses against herbivory suggest that the genetic mechanisms underlying these combinations of traits involve multiple genes and differential selection paradigms.

Convergent shifts in floral features resulting in observed shifts in pollinators are prevalent in the literature, including changes in corolla shape in Iochrominae (Solanaceae; Smith and Kriebel, 2018) and changes in stamen and stigma shape in *Salvia* L. (Lamiaceae; Kriebel et al., 2020, 2021). While many studies have investigated discrete shifts in flower color in relation to pollinators (Smith et al., 2008; McCarthy et al., 2015; Roguz et al., 2020; Garcia et al., 2021), evidence of significant pollinator-mediated selection on continuously varying flower color is lacking (Trunschke et al., 2021). Fewer studies have looked at the molecular evolution of genes responsible for pigment production, including anthocyanins

(Lu and Rausher, 2003; Ho and Smith, 2016; Wheeler et al., 2021) or carotenoids (Livingstone and Anderson, 2009; Cloutault et al., 2012; Cao et al., 2021). Ho and Smith (2016) found evidence of purifying selection on *CHI*, *F3H*, and *DFR*, providing additional evidence that gene expression may be more important than structural changes in the evolution of floral color, at least in the realm of losses of pigmentation. Across our sampling, we see evidence of positive selection on several anthocyanin producing genes including *FLS*, *ANS*, and *CHS* (Supplementary Table 6). Most shifts in flower color involve transitions from red and to green bracts and functional coloration, with a lack of transitions involving a loss of pigmentation (Figure 3). Even with the presence of flower color shifts between pollination syndromes, some anthocyanin genes show purifying selection while others in the pathway show positive selection across the phylogeny.

Our diversification analysis indicated no significant association between labellum shape and diversification (Supplementary Table 5) with the CID4 model of HiSSE showing the smallest AIC value. On one hand the lack of association is somewhat surprising since other studies have shown a strong association with diversification and traits important for pollination such as with floral pigments (Smith et al., 2010; Smith and Goldberg, 2015; Landis et al., 2018) and pollination trapping (Xue et al., 2020), corolla length (Moein et al., 2021), and corolla shape (Fernandez-Mazuecos et al., 2013). However, other traits important for pollination appear to be decoupled with diversification such as the lever mechanism in *Salvia* (Moein et al., 2021). Pollinator-mediated selection likely acts on more than one trait at a time to impact diversification, as observed by O'Meara et al. (2016), who found that the presence of a corolla, bilateral symmetry, and reduced stamen number synergistically act as key innovations to double diversification rates. The lack of impact of labellum shape on diversification may be due to the scope of the phylogeny, which was primarily focused on Neotropical *Costus*, and the rapid radiation of the group. Similar issues were previously seen in the study of nectar spurs in the *Pedilanthus* clade of *Euphorbia* (Cacho et al., 2010). If the phylogeny was expanded beyond the Neotropical Costaceae, the role of the labellum (and specifically labellum shape) might be considered a key innovation at the level of the family or even the order Zingiberales (Specht et al., 2012), similar to the labellum and pollinia in Orchidaceae (Zhang et al., 2017) and nectar spurs (Hodges and Arnold, 1995). However, formal testing of whether the labellum is a key innovation in Costaceae is still required.

From the selection analyses, we see that loci used for phylogenetic inference from different pipelines show different levels of selection (Figure 6). While neutral evolution of these markers might be advantageous for phylogenetic inference, many show signatures of purifying selection, lowering the overall rate of evolution (Edwards, 2009). In most cases this should not lead to biased results, though issues can

arise when investigating rapid radiations due to a lack of phylogenetic signal (Morrison et al., 2004; Grummer et al., 2018; Thomas et al., 2021). Additionally, many captured genes that regulate functional traits associated with floral phenotypes show evidence of non-neutral evolution including genes associated with floral fusion and the anthocyanin pathway. Genes evolving under positive selection do not predict congruence with other markers and therefore do not necessitate exclusion from phylogenetic inference (Roje, 2014). In fact, a phylotranscriptomic study of wild tomatoes (*Solanum* sect. *Lycopersicon*) showed that 3.8% of loci showed evidence of positive selection (Pease et al., 2016). From the data presented here, we see 13.3% (39 out of 293) of the loci used for phylogenetic inference showing positive selection (Supplementary Table 6). Notably, only 293 out of 957 loci that were used for phylogenetic inference were used for PAML analyses; specifically the loci remaining after our stringent filtering parameters. Therefore, the overall proportion of loci showing positive selection is likely much lower. Moving beyond a targeted capture approach to a fully genome wide data set may be necessary to fully resolve the evolutionary history of *Costus* and the Zingiberales overall. In fact, the genomic resources for the order are still quite limited, with one chromosome scale genome assembly in *Zingiber* Mill. (Li et al., 2021), several genomes for *Musa* L. (D'Hont et al., 2012; Droc et al., 2013; Wang et al., 2019), and three draft genomes of *Costus* (Grunder, 2021).

Conclusion

In summary we found that the recent radiation of Neotropical *Costus* encompasses unacknowledged diversity, and that the evolution of floral traits across the genus illustrates that several morphological traits associated with pollination syndromes and herbivory are correlated and are under selective pressure indicative of responses to involvement in complex ecological networks. Furthermore, selection detected in genomic regions provides evidence of the genetic underpinnings of these pollination-associated traits. The evidence presented here will guide many hypotheses, most likely to be tested at the population or species-pairs level, addressing the genomic mechanisms and the evolutionary ecology of these traits associated with the morphological diversity and evolution of pollination syndromes across this diverse lineage of the Neotropical Costaceae.

Data availability statement

The datasets and scripts generated for this study can be found in the Open Science Framework

<https://osf.io/fkj2x> and raw reads in NCBI BioProject <http://www.ncbi.nlm.nih.gov/bioproject/639561>. Raw sequencing data for whole genome sequencing and transcriptome sequencing, including Illumina and Nanopore data, are available with the following SRA numbers: SRR18516534 and SRR18516558 (numbers for specific accessions can be found in **Supplementary Table 3**). Both versions of the assembled *Costus bracteatus* genome are available on CoGe (<https://genomevolution.org/coge/>) with genome IDs 63651 (polished Flye assembly) and 63654 (polished MaSuRCA assembly).

Author contributions

CDS and AA conceived of the project and gathered the preliminary data. PM, HM-K, TA, and DS provided cultivated and field-collected materials of otherwise impossible-to-get taxa representing documented morphologic and biogeographic variation. EV and JL collected and analyzed the data and wrote the manuscript. CDS, DS, PM, HM-K, TA, and EV contributed to tissue collection, sampling, and database management. MP-V, AA, HP, NG, and CG collected the data and contributed to database management. CS collected and analyzed phylogenetic data. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by funds from Cornell University's College of Agriculture and Life Sciences and the School of Integrative Plant Science. No federal support was used for this research.

Acknowledgments

We thank the BTI Computational Biology Center for computational resources for analyses; additional resources were also provided by XSEDE through award DEB190021 awarded to JL, EV, AA, and CDS. We also thank Alexander Chautá, Shayla Salzman, and the Specht Lab for discussions and ideas about the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.874322/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Maximum likelihood (IQtree) concatenation phylogeny of 945 loci. The phylogeny is rooted with African *Costus* and Neotropical Costaceae accessions. Samples with wg in the ID indicates that the accession was included by assembling the target loci from whole genome sequencing data.

SUPPLEMENTARY FIGURE 2

Topologies obtained with concatenation vs. coalescent-based species tree analyses, showing just one node difference between the two. Branch lengths are proportional to the number of substitutions for the IQtree result and to coalescent units in the ASTRAL result.

SUPPLEMENTARY FIGURE 3

Species tree reconstruction by ASTRAL with local posterior probabilities above the branches. Pie charts illustrate the quartet scores for each node for the 945 loci, with red representing the current topology, blue the second most favored topology, and white the remaining one. Equal branch lengths were used to allow the reader to distinguish support values.

SUPPLEMENTARY FIGURE 4

Assembly stats for the (a) *Costus spicatus* transcriptome assembly and (b) genome assembly of *Costus bracteatus*. The N50 of each assembly is denoted with a blue line. Number of BUSCO genes from the embryophyta library (1,375 total genes) is shown for complete and single copy genes, complete and duplicate genes, fragmented, and missing genes.

SUPPLEMENTARY FIGURE 5

Branch specific omega values for (a) *PETAL LOSS (PTL)* and (b) *SUPERMAN* across the gene tree of 20 accessions with whole genome Illumina sequencing. Omega values representing purifying selection (<1) are in blue, values close to neutral evolution (~1) are in purple, and values of positive selection (>1) are shades of red. Ornithophilous and mellitophilous syndromes are indicated with a bee or hummingbird icon.

SUPPLEMENTARY FIGURE 6

Branch specific omega values for (a) *CHS4* and (b) *CHSY* across the gene tree of 20 accessions with whole genome Illumina sequencing. Omega values representing purifying selection (<1) are in blue, values close to neutral evolution (~1) are in purple, and values of positive selection (>1) are shades of red. Ornithophilous and mellitophilous syndromes are indicated with a bee or hummingbird icon.

SUPPLEMENTARY FIGURE 7

NeighborNetwork of the 20 species with Illumina genome sequencing highlighting conflicting splits. Accessions are coded as bee pollinated (blue), bird pollinated (red), or known hybrids (black).

SUPPLEMENTARY TABLE 1

Genus, species, identifiers, collector, number, and natural history collection for all the accessions included.

SUPPLEMENTARY TABLE 2

Binary traits scored for each accession including labellum shape (tubular/spreading), bract color (red/non-red), bract appendages (present/absent), honey mark (present/absent), labellum stripes (present/absent), functional color (red/non-red), callus (present/absent), and inflorescence indument (present/absent).

SUPPLEMENTARY TABLE 3

Genomic sequencing information for 20 accessions including the accession number, number of filtered SNPs retained, SRA number for Illumina sequencing, and SRA for Oxford Nanopore.

SUPPLEMENTARY TABLE 4

Phylogenetic D statistic estimates and p -values for the tests of phylogenetic D being significantly different from 0 (as expected under a Brownian threshold model) and 1 (random) for the analyzed traits.

Additionally, the Chi-squared estimate and associated p -value of the likelihood ratio test comparing ER vs. ARD models fitted for each trait are included.

SUPPLEMENTARY TABLE 5

Log-likelihood values (lnL), Akaike Information Criterion (AIC), and corrected AIC (AICc) for the different models of diversification fitted to the phylogeny of Neotropical *Costus*. The binary state speciation and extinction model (BISSE) as well as the hidden state speciation and extinction model (HiSSE) were fitted with the labellum shape character. The constant rate and additional hidden state models (CID2 and CID4) are independent from the labellum shape.

SUPPLEMENTARY TABLE 6

Omega values for each locus used in PAML, including which set that locus belongs to in the strip plot summary.

References

- Abrahamczyk, S., Weigend, M., Becker, K., Dannenberg, L. S., Eberz, J., Atella-Hödtke, N., et al. (2022). Influence of plant reproductive systems on the evolution of hummingbird pollination. *Ecol. Evol.* 12:e8621. doi: 10.1002/ece3.8621
- Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M. (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* 9, 841–857. doi: 10.1105/tpc.9.6.841
- Almeida, A. M. R., Yockteng, R., Otoni, W. C., and Specht, C. D. (2015). Positive selection on the K domain of the AGAMOUS protein in the Zingiberales suggests a mechanism for the evolution of androecial morphology. *Evodevo* 6:7. doi: 10.1186/s13227-015-0002-x
- Almeida, A. M. R., Yockteng, R., Schnable, J., Alvarez-Buylla, E. R., Freeling, M., and Specht, C. D. (2014). Co-option of the polarity gene network shapes filament morphology in angiosperms. *Sci. Rep.* 4:6194. doi: 10.1038/srep06194
- Almeida, J., Rocheta, M., and Galego, L. (1997). Genetic control of flower shape in *Antirrhinum majus*. *Development* 124, 1387–1392. doi: 10.1242/dev.124.7.1387
- André, T., Salzman, S., Wendt, T., and Specht, C. (2016). Speciation dynamics and biogeography of Neotropical spiral gingers (Costaceae). *Mol. Phylogenet. Evol.* 103, 55–63. doi: 10.1016/j.ympev.2016.07.008
- Ashworth, L., Aguilar, R., Martín-Rodríguez, S., Lopezaraiza-Mikel, M., Avila-Sakar, G., Rosas-Guerrero, V., et al. (2015). *Pollination Syndromes: a Global Pattern of Convergent Evolution Driven by the Most Effective Pollinator*. Berlin: Springer International Publishing, 203–224. doi: 10.1007/978-3-319-19932-0_11
- Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., and Marth, G. T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* 27, 1691–1692. doi: 10.1093/bioinformatics/btr174
- Barrio, R. Á., Hernández-Machado, A., Varea, C., Romero-Arias, J. R., and Alvarez-Buylla, E. (2010). Flower development as an interplay between dynamical physical fields and genetic networks. *PLoS One* 5:e13523. doi: 10.1371/journal.pone.0013523
- Bartlett, M. E., and Specht, C. D. (2010). Evidence for the involvement of Globosa-like gene duplications and expression divergence in the evolution of floral morphology in the Zingiberales. *New Phytol.* 187, 521–541. doi: 10.1111/j.1469-8137.2010.03279.x
- Bawa, K. S. (1990). Plant-pollinator interactions in tropical rain forests. *Annu. Rev. Ecol. Syst.* 21, 399–422. doi: 10.1146/annurev.es.21.110190.002151
- Beaulieu, J. M., and O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* 65, 583–601. doi: 10.1093/sysbio/syw022
- Bergamo, P. J., Rech, A. R., Brito, V. L. G., and Sazima, M. (2016). Flower colour and visitation rates of *Costus arabicus* support the “bee avoidance” hypothesis for red-reflecting hummingbird-pollinated flowers. *Funct. Ecol.* 30, 710–720. doi: 10.1111/1365-2435.12537
- Bonferroni, C. E. (1935). “Il calcolo delle assicurazioni su gruppi di teste,” in *Studi in Onore del Professore Salvatore Ortu Carboni*, ed. S. O. Carboni (Rome: Bardi), 13–60.
- Boyden, G. S., Donoghue, M. J., and Howarth, D. G. (2012). Duplications and expression of RADIALIS-like genes in Dipsacales. *Int. J. Plant Sci.* 173, 971–983. doi: 10.1086/667626
- Broad Institute (2019). *Picard Toolkit*. Cambridge, MA: Broad Institute
- Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., et al. (2017). A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* 18, 762–776. doi: 10.1016/j.celrep.2016.12.063
- Cabin, Z., Derieg, N. J., Garton, A., Ngo, T., Quezada, A., Gasseholm, C., et al. (2022). Non-pollinator selection for a floral homeotic mutant conferring loss of nectar reward in *Aquilegia coerulea*. *Curr. Biol.* 32, 1332–1341.e5. doi: 10.1016/j.cub.2022.01.066
- Cacho, N. I., Berry, P. E., Olson, M. E., Steinmann, V. W., and Baum, D. A. (2010). Are spurred cyathia a key innovation? Molecular systematics and trait evolution in the slipper spurge (Pedilanthus clade: euphorbia, Euphorbiaceae). *Am. J. Bot.* 97, 493–510. doi: 10.3732/ajb.0900090
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421
- Cao, W., Wang, P., Yang, L., Fang, Z., Zhang, Y., Zhuang, M., et al. (2021). Carotenoid biosynthetic genes in cabbage: genome-wide identification, evolution, and expression analysis. *Genes* 12:2027. doi: 10.3390/genes12122027
- Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Carlsen, M. M., Fér, T., Schmickl, R., Leong-Škorničková, J., Newman, M., and Kress, W. J. (2018). Resolving the rapid plant radiation of early diverging lineages in the tropical Zingiberales: pushing the limits of genomic data. *Mol. Phylogenet. Evol.* 128, 55–68. doi: 10.1016/j.ympev.2018.07.020
- Caruso, C. M., Eisen, K. E., Martin, R. A., and Sletvold, N. (2019). A meta-analysis of the agents of selection on floral traits. *Evolution* 73, 4–14.
- Castellanos, M. C., Wilson, P., and Thomson, J. D. (2004). “Anti-bee” and “pro-bird” changes during the evolution of hummingbird pollination in *Penstemon* flowers. *J. Evol. Biol.* 17, 876–885. doi: 10.1111/j.1420-9101.2004.00729.x
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Cloutalt, J., Peltier, D., Soufflet-Freslon, V., Briard, M., and Geoffriau, E. (2012). Differential selection on carotenoid biosynthesis genes as a function of gene position in the metabolic pathway: a study on the carrot and dicots. *PLoS One* 7:e38724. doi: 10.1371/journal.pone.0038724
- Coen, E. S., Nugent, J. M., Luo, D., Bradley, D., Cubas, P., Chadwick, M., et al. (1995). Evolution of floral symmetry. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 350, 35–38. doi: 10.1098/rstb.1995.0134
- Coimbra, G., Araujo, C., Bergamo, P. J., Freitas, L., and Rodríguez-Gironés, M. A. (2020). Flower conspicuousness to bees across pollination systems: a generalized test of the bee-avoidance hypothesis. *Front. Plant Sci.* 11:558684. doi: 10.3389/fpls.2020.558684
- D'Hont, A., Denoeud, F., Aury, J.-M., Baurens, F.-C., Carreel, F., Garsmeur, O., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488, 213–217. doi: 10.1038/nature11241

- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Darwin, C. (1892). *Various contrivances by which orchids are fertilised by insects*. 2d Edn. New York, NY: Appleton.
- De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. doi: 10.1093/bioinformatics/bty149
- Dellinger, A. S. (2020). Pollination syndromes in the 21st century: where do we stand and where may we go? *New Phytol.* 228, 1193–1213. doi: 10.1111/nph.16793
- Dellinger, A. S., Paun, O., Baar, J., Tensch, E. M., Fernández-Fernández, D., and Schönenberger, J. (2022). Population structure in Neotropical plants: Integrating pollination biology, topography and climatic niches. *Mol. Ecol.* 31, 2264–2280. doi: 10.1111/mec.16403
- Disch, S., Anastasiou, E., Sharma, V. K., Laux, T., Fletcher, J. C., and Lenhard, M. (2006). The E3 ubiquitin ligase BIG BROTHER controls *Arabidopsis* organ size in a dosage-dependent manner. *Curr. Biol.* 16, 272–279. doi: 10.1016/j.cub.2005.12.026
- Droc, G., Larivière, D., Guignon, V., Yahiaoui, N., This, D., Garsmeur, O., et al. (2013). The Banana Genome Hub. doi: 10.1093/database/bat035
- Durán-Castillo, M., Hudson, A., Wilson, Y., Field, D. L., and Twyford, A. D. (2022). A phylogeny of *Antirrhinum* reveals parallel evolution of alpine morphology. *New Phytol.* 233, 1426–1439. doi: 10.1111/nph.17581
- Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19:1349. doi: 10.1093/nar/19.6.1349
- Edwards, S. V. (2009). Natural selection and phylogenetic analysis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8799–8800. doi: 10.1073/pnas.0904103106
- Ellis, C. M., Nagpal, P., Young, J. C., Hagen, G., Guilfoyle, T. J., and Reed, J. W. (2005). AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development* 132, 4563–4574. doi: 10.1242/dev.02012
- Faegri, K., and Pijl, L. (1979). *The Principles Of Pollination Ecology*. Amsterdam: Elsevier, 244–244.
- Faircloth, B. C. (2016). PHYLUCES is a software package for the analysis of conserved genomic loci. *Bioinformatics* 32, 786–788. doi: 10.1093/bioinformatics/btv646
- Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., and Glenn, T. C. (2012). Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales. *Syst. Biol.* 61, 717–726. doi: 10.1093/sysbio/sys004
- Fenster, C. B., Armbruster, W. S., Wilson, P., Dudash, M. R., and Thomson, J. D. (2004). Pollination syndromes and floral specialization. *Annu. Rev. Ecol. Syst.* 35, 375–403. doi: 10.1146/annurev.ecolsys.34.011802.132347
- Fernandez-Mazuecos, M., Blanco-Pastor, J. L., Gomez, J. M., and Vargas, P. (2013). Corolla morphology influences diversification rates in bifid toadflaxes (*Linaria* sect. *Versicolores*). *Ann. Bot.* 112, 1705–1722. doi: 10.1093/aob/mct214
- Frachon, L., and Schiestl, F. P. (2021). Rapid genomic evolution in *Brassica rapa* with bumblebee selection in experimental evolution. *bioRxiv* [Preprint]. doi: 10.1101/2021.12.02.470896
- Fragoso-Martínez, I., Martínez-Gordillo, M., Salazar, G. A., Sazatornil, F., Jenks, A. A., García Peña, M., et al. (2018). Phylogeny of the Neotropical sages (*Salvia* subg. *Calosphaea*; *Lamiaceae*) and insights into pollinator and area shifts. *Plant Syst. Evol.* 304, 43–55. doi: 10.1007/s00606-017-1445-4
- Freitas, L., Galetto, L., and Sazima, M. (2006). Pollination by hummingbirds and bees in eight syntopic species and a putative hybrid of Ericaceae in Southeastern Brazil. *Osterr. bot. Z.* 258, 49–61. doi: 10.1007/s00606-005-0392-7
- Fritz, S. A., and Purvis, A. (2010). Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conserv. Biol.* 24, 1042–1051. doi: 10.1111/j.1523-1739.2010.01455.x
- García, J. E., Hannah, L., Shrestha, M., Burd, M., and Dyer, A. G. (2021). Fly pollination drives convergence of flower coloration. *New Phytol.* 233, 52–61. doi: 10.1111/nph.17696
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Grubbaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., et al. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* 20:8. doi: 10.1186/s13059-018-1618-7
- Grummer, J. A., Morando, M. M., Avila, L. J., Sites, J. W. Jr., and Leaché, A. D. (2018). Phylogenomic evidence for a recent and rapid radiation of lizards in the Patagonian *Liolaemus fitzingerii* species group. *Mol. Phylogenet. Evol.* 125, 243–254. doi: 10.1016/j.ympev.2018.03.023
- Grunder, N. (2021). *Genome Consequences of Hybridization in Neotropical Costus*. [Ph.D thesis] Long Beach, CA: California State University
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Ha, C. M., Jun, J. H., and Fletcher, J. C. (2010). Control of *Arabidopsis* leaf morphogenesis through regulation of the YABBY and KNOX families of transcription factors. *Genetics* 186, 197–206. doi: 10.1534/genetics.110.118703
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084
- Herzog, M., Dorne, A. M., and Grellet, F. (1995). GASA, a gibberellin-regulated gene family from *Arabidopsis thaliana* related to the tomato GAST1 gene. *Plant Mol. Biol.* 27, 743–752. doi: 10.1007/bf00020227
- Ho, W. W., and Smith, S. D. (2016). Molecular evolution of anthocyanin pigmentation genes following losses of flower color. *BMC Evol. Biol.* 16:98. doi: 10.1186/s12862-016-0675-3
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281
- Hodges, S. A., and Arnold, M. L. (1995). Spurring plant diversification: are floral nectar spurs a key innovation? *Proc. R. Soc. Lond. B* 262, 343–348. doi: 10.1098/rspb.1995.0215
- Holt, J., and McMillan, L. (2014). Merging of multi-string BWTs with applications. *Bioinformatics* 30, 3524–3531. doi: 10.1093/bioinformatics/btu584
- Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.1105/tpc.7.7.1071
- Huelsbeck, J. P., Nielsen, R., and Bollback, J. P. (2003). Stochastic mapping of morphological characters. *Syst. Biol.* 52, 131–158. doi: 10.1080/10635150390192780
- Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Irwin, R. E., Adler, L. S., and Brody, A. K. (2004). The dual role of floral traits: pollinator attraction and plant defense. *Ecology* 85, 1503–1511. doi: 10.1890/03-0390
- Jeffares, D. C., Tomiczek, B., Sojo, V., and dos Reis, M. (2015). A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. *Methods Mol. Biol.* 1201, 65–90. doi: 10.1007/978-1-4939-1438-8_4
- Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., et al. (2016). HybPiper: extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl. Plant Sci.* 4, 1600016–1600016. doi: 10.3732/apps.1600016
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., and Jermin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Katoh, K., and Toh, H. (2010). Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* 26, 1899–1900. doi: 10.1093/bioinformatics/btq224
- Kay, K. M., and Schemske, D. W. (2003). Pollinator assemblages and visitation rates for 11 species of neotropical *Costus* (Costaceae). *Biotropica* 35, 198–207. doi: 10.1111/j.1744-7429.2003.tb00279.x
- Kay, K. M., Reeves, P. A., Olmstead, R. G., and Schemske, D. W. (2005). Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. *Am. J. Bot.* 92, 1899–1910. doi: 10.3732/ajb.92.11.1899
- Kim, J., and Sanderson, M. J. (2008). Penalized likelihood phylogenetic inference: bridging the parsimony-likelihood gap. *Syst. Biol.* 57, 665–674. doi: 10.1080/10635150802422274
- Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi: 10.1038/s41587-019-0072-8

- Konieczny, A., and Ausubel, F. M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* 4, 403–410. doi: 10.1046/j.1365-313X.1993.04020403.x
- Koo, S. C., Bracko, O., Park, M. S., Schwab, R., Chun, H. J., Park, K. M., et al. (2010). Control of lateral organ development and flowering time by the *Arabidopsis thaliana* MADS-box gene AGAMOUS-LIKE6. *Plant J.* 62, 807–816. doi: 10.1111/j.1365-313X.2010.04192.x
- Kotilainen, M., Helariutta, Y., Mehto, M., Pollanen, E., Albert, V. A., Elomaa, P., et al. (1999). GEG participates in the regulation of cell and organ shape during corolla and carpel development in *Gerbera hybrida*. *Plant Cell* 11, 1093–1104. doi: 10.2307/3870801
- Kriebel, R., Drew, B. T., González-Gallegos, J. G., Celep, F., Antar, G. M., Pastore, J. F. B., et al. (2021). Stigma shape shifting in sages (*Salvia*: lamiaceae): hummingbirds guided the evolution of New World floral features. *Bot. J. Linn. Soc.* 199, 428–448. doi: 10.1093/botlinnean/boab096
- Kriebel, R., Drew, B., González-Gallegos, J. G., Celep, F., Heeg, L., Mahdjoub, M. M., et al. (2020). Pollinator shifts, contingent evolution, and evolutionary constraint drive floral disparity in *Salvia* (Lamiaceae): evidence from morphometrics and phylogenetic comparative methods. *Evolution* 74, 1335–1355. doi: 10.1111/evo.14030
- Krizek, B. A., Lewis, M. W., and Fletcher, J. C. (2006). RABBIT EARS is a second-whorl repressor of AGAMOUS that maintains spatial boundaries in *Arabidopsis* flowers. *Plant J.* 45, 369–383. doi: 10.1111/j.1365-313X.2005.02633.x
- Lagomarsino, L. P., Condamine, F. L., Antonelli, A., Mulch, A., and Davis, C. C. (2016). The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* 210, 1430–1442. doi: 10.1111/nph.13920
- Lampugnani, E. R., Kilinc, A., and Smyth, D. R. (2012). PETAL LOSS is a boundary gene that inhibits growth between developing sepals in *Arabidopsis thaliana*. *Plant J.* 71, 724–735. doi: 10.1111/j.1365-313X.2012.05023.x
- Landis, J. B., Bell, C. D., Hernandez, M., Zenil-Ferguson, R., McCarthy, E. W., Soltis, D. E., et al. (2018). Evolution of floral traits and impact of reproductive mode on diversification in the phlox family (Polemoniaceae). *Mol. Phylogenet. Evol.* 127, 878–890. doi: 10.1016/j.ympev.2018.06.035
- Landis, J. B., Kurti, A., Lawhorn, A. J., Litt, A., and McCarthy, E. W. (2020). Differential gene expression with an emphasis on floral organ size differences in natural and synthetic polyploids of *Nicotiana tabacum* (Solanaceae). *Genes* 11:1097. doi: 10.3390/genes11091097
- Lanfear, R., Calcott, B., Kainer, D., Mayer, C., and Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* 14:82–82. doi: 10.1186/1471-2148-14-82
- Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531
- Lee, Y. K., Kim, G.-T., Kim, I.-J., Park, J., Kwak, S.-S., Choi, G., et al. (2006). LONGIFOLIA1 and LONGIFOLIA2, two homologous genes, regulate longitudinal cell elongation in *Arabidopsis*. *Development* 133, 4305–4314. doi: 10.1242/dev.02604
- Lex, A., Gehlenborg, N., Strobel, H., Vuilleumot, R., and Pfister, H. (2014). UpSet: visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* 20, 1983–1992. doi: 10.1109/TVCG.2014.2346248
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, H.-L., Wu, L., Dong, Z., Jiang, Y., Jiang, S., Xing, H., et al. (2021). Haplotype-resolved genome of diploid ginger (*Zingiber officinale*) and its unique gingerol biosynthetic pathway. *Hortic Res.* 8:189. doi: 10.1038/s41438-021-00627-7
- Livingstone, K., and Anderson, S. (2009). Patterns of variation in the evolution of carotenoid biosynthetic pathway enzymes of higher plants. *J. Hered.* 100, 754–761. doi: 10.1093/jhered/esp026
- Lu, Y., and Rausher, M. D. (2003). Evolutionary rate variation in anthocyanin pathway genes. *Mol. Biol. Evol.* 20, 1844–1853. doi: 10.1093/molbev/msg197
- Lunau, K. (2014). Visual ecology of flies with particular reference to colour vision and colour preferences. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 200, 497–512. doi: 10.1007/s00359-014-0895-1
- Lunau, K., and Maier, E. J. (1995). Innate colour preferences of flower visitors. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 177, 1–19. doi: 10.1007/bf00243394
- Lunau, K., Papiorek, S., Eltz, T., and Sazima, M. (2011). Avoidance of achromatic colours by bees provides a private niche for hummingbirds. *J. Exp. Biol.* 214, 1607–1612. doi: 10.1242/jeb.052688
- Maas, P. J. M. (1972). Costoideae (Zingiberaceae). *Flora Neotropica* 8, 1–139. doi: 10.1093/aob/mch177
- Maas, P. J. M. (1977). *Renalmia* (Zingiberaceae-Zingiberoideae) Costoideae (Additions)(Zingiberaceae). *Flora Neotropica* 18, 1–218.
- Maas-van de Kamer, H., Maas, P. J. M., Wieringa, J. J., and Specht, C. D. (2016). Monograph of African Costaceae. *Blumea* 61, 280–318. doi: 10.3767/000651916X694445
- Maddison, W. P., Midford, P. E., and Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* 56, 701–710. doi: 10.1080/10635150701607033
- Mai, U., and Mirarab, S. (2018). TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics* 19:272–272. doi: 10.1186/s12864-018-4620-2
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200
- McCarthy, E. W., Arnold, S. E. J., Chittka, L., Le Comber, S. C., Verity, R., Dodsworth, S., et al. (2015). The effect of polyploidy and hybridization on the evolution of floral colour in *Nicotiana* (Solanaceae). *Ann. Bot.* 115, 1117–1131. doi: 10.1093/aob/mcv048
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- Mendoza, L., Thieffry, D., and Alvarez-Buylla, E. R. (1999). Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics* 15, 593–606. doi: 10.1093/bioinformatics/15.7.593
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Moein, F., Jamzad, Z., Rahiminejad, M., Landis, J. B., Mirtadzadini, M., Soltis, D. E., et al. (2021). Towards a global perspective for *Salvia* L.: phylogeny, diversification, and floral evolution. *bioRxiv* [Preprint]. doi: 10.1101/2021.12.16.473009
- Morioka, K., Yockteng, R., Almeida, A. M. R., and Specht, C. D. (2015). Loss of YABBY2-Like gene expression may underlie the evolution of the laminar style in *Canna* and contribute to floral morphological diversity in the Zingiberales. *Front. Plant Sci.* 6:1106. doi: 10.3389/fpls.2015.01106
- Morrell, P. L., Toleno, D. M., Lundy, K. E., and Clegg, M. T. (2005). Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite high rates of self-fertilization. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2442–2447. doi: 10.1073/pnas.0409804102
- Morrison, C. L., Rios, R., and Duffy, J. E. (2004). Phylogenetic evidence for an ancient rapid radiation of Caribbean sponge-dwelling snapping shrimps (Synalpheus). *Mol. Phylogenet. Evol.* 30, 563–581. doi: 10.1016/S1055-7903(03)00252-5
- Niu, S., Song, Q., Koiwa, H., Qiao, D., Zhao, D., Chen, Z., et al. (2019). Genetic diversity, linkage disequilibrium, and population structure analysis of the tea plant (*Camellia sinensis*) from an origin center, Guizhou plateau, using genome-wide SNPs developed by genotyping-by-sequencing. *BMC Plant Biol.* 19:328. doi: 10.1186/s12870-019-1917-5
- Nogueira, R. R., Santos, D. F. B., Calixto, E. S., Torezan-Silingardi, H. M., and Del-Claro, K. (2021). Negative effects of ant-plant interaction on pollination: costs of a mutualism. *Sociobiology* 68:e7259. doi: 10.13102/sociobiology.v68i4.7259
- O'Meara, B. C., Smith, S. D., Scott Armbruster, W., Harder, L. D., Hardy, C. R., Hileman, L. C., et al. (2016). Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proc. R. Soc. B* 283:20152304. doi: 10.1098/rspb.2015.2304
- Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., et al. (2013). The caper package: comparative analysis of phylogenetics and evolution in R. *R package version* 5, 1–36.
- Ortiz, E. M. (2019). Vcf2phylyp V2.0: Convert A Vcf Matrix Into Several Matrix Formats For Phylogenetic Analysis. doi: 10.5281/zenodo.2540861
- Pagel, M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. Series B* 255, 37–45. doi: 10.1098/rspb.1994.0006

- Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633
- Pease, J. B., Haak, D. C., Hahn, M. W., and Moyle, L. C. (2016). Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biol.* 14:e1002379. doi: 10.1371/journal.pbio.1002379
- Pérez-Escobar, O. A., Dodsworth, S., Bogarín, D., Bellot, S., Balbuena, J. A., Schley, R. J., et al. (2021). Hundreds of nuclear and plastid loci yield novel insights into orchid relationships. *Am. J. Bot.* 108, 1166–1180. doi: 10.1002/ajb2.1702
- Portik, D. M., Smith, L. L., and Bi, K. (2016). An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). *Mol. Ecol. Resour.* 16, 1069–1083. doi: 10.1111/1755-0998.12541
- Pringle, E. G., Dirzo, R., and Gordon, D. M. (2012). Plant defense, herbivory, and the growth of *Cordia alliodora* trees and their symbiotic Azteca ant colonies. *Oecologia* 170, 677–685. doi: 10.1007/s00442-012-2340-x
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Ranwez, V., Harispe, S., Delsuc, F., and Douzery, E. J. P. (2011). MACSE: multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PLoS One* 6:e22594. doi: 10.1371/journal.pone.0022594
- Raven, P. H. (1972). Why are bird-visited flowers predominantly red? *Evolution* 26:674. doi: 10.1111/j.1558-5646.1972.tb01975.x
- Renner, S. S., and Feil, J. P. (1993). Pollinators of tropical dioecious angiosperms. *Am. J. Bot.* 80, 1100–1107. doi: 10.1002/j.1537-2197.1993.tb15337.x
- Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x
- Roda, F., Liu, H., Wilkinson, M. J., Walter, G. M., James, M. E., Bernal, D. M., et al. (2013). Convergence and divergence during the adaptation to similar environments by an Australian groundsel. *Evolution* 67, 2515–2529. doi: 10.1111/evo.12136
- Rodríguez-Gironés, M. A., and Santamaría, L. (2004). Why are so many bird flowers red? *PLoS Biol.* 2:e350. doi: 10.1371/journal.pbio.0020350
- Roguz, K., Gallagher, M. K., Senden, E., Bar-Lev, Y., Lebel, M., Heliczler, R., et al. (2020). All the colors of the rainbow: diversification of flower color and intraspecific color variation in the genus *Iris*. *Front. Plant Sci.* 11:1519. doi: 10.3389/fpls.2020.569811
- Roje, D. M. (2014). Evaluating the effects of non-neutral molecular markers on phylogeny inference. *PLoS One* 9:e87428. doi: 10.1371/journal.pone.0087428
- Rosas-Guerrero, V., Aguilar, R., Martín-Rodríguez, S., Ashworth, L., Lopezzariza-Mikel, M., Bastida, J. M., et al. (2014). A quantitative review of pollination syndromes: do floral traits predict effective pollinators? *Ecol. Lett.* 17, 388–400. doi: 10.1111/ele.12224
- Rosumek, F. B., Silveira, F. A. O., de, S., Neves, F., de, U., Barbosa, N. P., et al. (2009). Ants on plants: a meta-analysis of the role of ants as plant biotic defenses. *Oecologia* 160, 537–549. doi: 10.1007/s00442-009-1309-x
- Rowan, B. A., Seymour, D. K., Chae, E., Lundberg, D. S., and Weigel, D. (2017). Methods for Genotyping-by-Sequencing. *Genotyping* 221–242. doi: 10.1007/978-1-4939-6442-0_16
- Sakai, H., Medrano, L. J., and Meyerowitz, E. M. (1995). Role of SUPERMAN in maintaining *Arabidopsis* floral whorl boundaries. *Nature* 378, 199–203. doi: 10.1038/378199a0
- Salzman, S., Driscoll, H. E., Renner, T., André, T., Shen, S., and Specht, C. D. (2015). Spiraling into history: a molecular phylogeny and investigation of biogeographic origins and floral evolution for the genus *Costus*. *Syst. Bot.* 40, 104–115. doi: 10.1600/036364415x686404
- Sass, C., Iles, W. J. D., Barrett, C. F., Smith, S. Y., and Specht, C. D. (2016). Revisiting the Zingiberales: using multiplexed exon capture to resolve ancient and recent phylogenetic splits in a charismatic plant lineage. *PeerJ* 4:e1584. doi: 10.7717/peerj.1584
- Schalamun, M., Nagar, R., Kainer, D., Beavan, E., Eccles, D., Rathjen, J. P., et al. (2019). Harnessing the MinION: an example of how to establish long-read sequencing in a laboratory using challenging plant tissue from *Eucalyptus pauciflora*. *Mol. Ecol. Resour.* 19, 77–89. doi: 10.1111/1755-0998.12938
- Schemske, D. W. (1980). The evolutionary significance of extrafloral nectar production by *Costus woodsonii* (Zingiberaceae): an experimental analysis of ant protection. *J. Ecol.* 68, 959–967. doi: 10.2307/2259468
- Schemske, D. W. (1982). Ecological correlates of a Neotropical mutualism: ant assemblages at *Costus* extrafloral nectaries. *Ecology* 63, 932–941. doi: 10.2307/1937233
- Schemske, D. W., and Horvitz, C. C. (1984). Variation among floral visitors in pollination ability: a precondition for mutualism specialization. *Science* 225, 519–521. doi: 10.1126/science.225.4661.519
- Schley, R. J., de la Estrella, M., Pérez-Escobar, O. A., Bruneau, A., Barraclough, T., Forest, F., et al. (2018). Is Amazonia a “museum” for Neotropical trees? The evolution of the *Brownea* clade (Detarioideae, Leguminosae). *Mol. Phylogenet. Evol.* 126, 279–292. doi: 10.1016/j.ympev.2018.04.029
- Schley, R. J., Twyford, A. D., and Toby Pennington, R. (2022). Hybridization: a “double-edged sword” for Neotropical plant diversity. *Bot. J. Linnean Soc.* 199, 331–356. doi: 10.1093/botlinnean/boab070
- Serrano-Serrano, M. L., Rolland, J., Clark, J. L., Salamin, N., and Perret, M. (2017). Hummingbird pollination and the diversification of angiosperms: an old and successful association in Gesneriaceae. *Proc. R. Soc. B* 284:20162816. doi: 10.1098/rspb.2016.2816
- Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11:e0163962. doi: 10.1371/journal.pone.0163962
- Shi, L., Gast, R. T., Gopalraj, M., and Olszewski, N. E. (1992). Characterization of a shoot-specific, GA3- and ABA-regulated gene from tomato. *Plant J.* 2, 153–159.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Smith, S. D., and Goldberg, E. E. (2015). Tempo and mode of flower color evolution. *Am. J. Bot.* 102, 1014–1025. doi: 10.3732/ajb.1500163
- Smith, S. D., and Kriebel, R. (2018). Convergent evolution of floral shape tied to pollinator shifts in Iochrominae (Solanaceae). *Evolution* 72, 688–697. doi: 10.1111/evo.13416
- Smith, S. D., Ané, C., and Baum, D. A. (2008). The role of pollinator shifts in the floral diversification of *Iochroma* (Solanaceae). *Evolution* 62, 793–806. doi: 10.1111/j.1558-5646.2008.00327.x
- Smith, S. D., Miller, R. E., Otto, S. P., FitzJohn, R. G., and Rausher, G. (2010). “The effects of flower color transitions on diversification rates in morning glories (*Ipomea* subg. *Quamoclit*, Convolvulaceae),” in *Darwin’s Heritage Today: Proceedings of the Darwin 200 Beijing International Conference*, ed. M. Long (Beijing: Higher Education Press).
- Specht, C. D. (2006). Systematics and evolution of the tropical monocot family Costaceae (Zingiberales): a multiple dataset approach. *Syst. Bot.* 31, 89–106.
- Specht, C. D., Kress, W. J., Stevenson, D. W., and DeSalle, R. (2001). A molecular phylogeny of Costaceae (Zingiberales). *Mol. Phylogenet. Evol.* 21, 333–345. doi: 10.1006/mpev.2001.1029
- Specht, C. D., Yockteng, R., Almeida, A. M., Kirchoff, B. K., and Kress, W. J. (2012). Homoplasy, pollination, and emerging complexity during the evolution of floral development in the tropical gingers (Zingiberales). *Bot. Rev.* 78, 440–462. doi: 10.1007/s12229-012-9111-6
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stebbins, G. L. (1970). Adaptive radiation of reproductive characteristics in angiosperms. I: pollination mechanisms. *Annu. Rev. Ecol. Syst.* 1, 307–326. doi: 10.1146/annurev.es.01.110170.001515
- Thomas, A. E., Igea, J., Meudt, H. M., Albach, D. C., Lee, W. G., and Tanentzap, A. J. (2021). Using target sequence capture to improve the phylogenetic resolution of a rapid radiation in New Zealand *Veronica*. *Am. J. Bot.* 108, 1289–1306. doi: 10.1002/ajb2.1678
- Thomson, J. (2003). When is it mutualism? (American Society of Naturalists Presidential Address, 2001). *Am. Nat.* 162, S1–S9. doi: 10.1086/378683
- Torres Jimenez, M. F., Stone, G. N., Sanchez, A., and Richardson, J. E. (2021). Comparative phylogeography of an ant-plant mutualism: an encounter in the Andes. *Glob. Planet. Change* 205:103598. doi: 10.1016/j.gloplacha.2021.103598
- Trunschke, J., Lunau, K., Pyke, G. H., Ren, Z.-X., and Wang, H. (2021). Flower color evolution and the evidence of pollinator-mediated selection. *Front. Plant Sci.* 12:617851. doi: 10.3389/fpls.2021.617851
- Valderrama, E., Richardson, J. E., Kidner, C. A., Madriñán, S., and Stone, G. N. (2018). Transcriptome mining for phylogenetic markers in a recently radiated genus of tropical plants (*Renalmia* L.f., Zingiberaceae). *Mol. Phylogenet. Evol.* 119, 13–24. doi: 10.1016/j.ympev.2017.10.001

- Valderrama, E., Sass, C., Pinilla-Vargas, M., Skinner, D., Maas, P. J. M., Maas-van, et al. (2020). Unraveling the spiraling radiation: a phylogenomic analysis of Neotropical *Costus* L. *Front. Plant Sci.* 11:1195. doi: 10.3389/fpls.2020.01195
- Van der Niet, T., Peakall, R., and Johnson, S. D. (2014). Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Ann. Bot.* 113, 199–211. doi: 10.1093/aob/mct290
- Vargas, O. M., Goldston, B., Gossenbacher, D. L., and Kay, K. M. (2020). Patterns of speciation are similar across mountainous and lowland regions for a Neotropical plant radiation (Costaceae: *Costus*). *Evolution* 74, 2644–2661. doi: 10.1111/evo.14108
- Vasimuddin, M., Misra, S., Li, H., and Aluru, S. (2019). “Efficient architecture-aware acceleration of BWA-MEM for multicore systems,” in *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. Piscataway, NJ: IEEE 314–324. doi: 10.1109/IPDPS.2019.00041
- Vurtture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., et al. (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* 33, 2202–2204. doi: 10.1093/bioinformatics/btx153
- Wang, J. R., Holt, J., McMillan, L., and Jones, C. D. (2018). FMLRC: hybrid long read error correction using an FM-index. *BMC Bioinform.* 19:50. doi: 10.1186/s12859-018-2051-3
- Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., et al. (2020). Treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. *Mol. Biol. Evol.* 37, 599–603. doi: 10.1093/molbev/msz240
- Wang, Z., Miao, H., Liu, J., Xu, B., Yao, X., Xu, C., et al. (2019). *Musa balbisiana* genome reveals subgenome evolution and functional divergence. *Nat. Plants* 5, 810–821. doi: 10.1038/s41477-019-0452-6
- Wheeler, L. C., Walker, J. F., Ng, J., Deanna, R., Dunbar-Wallis, A., Backes, A., et al. (2021). Transcription factors evolve faster than their structural gene targets in the flavonoid pigment pathway. *Mol. Biol. Evol.* 39:msac044. doi: 10.1093/molbev/msac044
- Whelan, R. J., Ayre, D. J., and Beynon, F. M. (2009). The birds and the bees: pollinator behaviour and variation in the mating system of the rare shrub *Grevillea macleayana*. *Ann. Bot.* 103, 1395–1401. doi: 10.1093/aob/mcp091
- Xue, B., Guo, X., Landis, J. B., Sun, M., Tang, C. C., Soltis, P. S., et al. (2020). Accelerated diversification correlated with functional traits shapes extant diversity of the early divergent angiosperm family Annonaceae. *Mol. Phylogenet. Evol.* 142:106659. doi: 10.1016/j.ympev.2019.106659
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Yockteng, R., Almeida, A. M. R., Yee, S., Andre, T., Hill, C., and Specht, C. D. (2013). A method for extracting high-quality RNA from diverse plants for next-generation sequencing and gene expression analyses. *Appl. Plant Sci.* 1:apps.1300070 doi: 10.3732/apps.1300070
- Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2017). Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi: 10.1111/2041-210x.12628
- Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018b). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* 19, 153. doi: 10.1186/s12859-018-2129-y
- Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., and Yang, T.-L. (2018a). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35, 1786–1788. doi: 10.1093/bioinformatics/bty875
- Zhang, G.-Q., Liu, K.-W., Li, Z., Lohaus, R., Hsiao, Y.-Y., Niu, S.-C., et al. (2017). The *Apostasia* genome and the evolution of orchids. *Nature* 549, 379–383. doi: 10.1038/nature23897
- Zhang, H., Deng, W., Lu, C., He, M., and Yan, H. (2022). SMRT sequencing of full-length transcriptome and gene expression analysis in two chemical types of *Pogostemon cablin* (Blanco) Benth. *PeerJ* 10:e12940. doi: 10.7717/peerj.12940
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28, 3326–3328. doi: 10.1093/bioinformatics/bts606
- Zhu, X., Dong, L., Jiang, L., Li, H., Sun, L., Zhang, H., et al. (2016). Constructing a linkage-linkage disequilibrium map using dominant-segregating markers. *DNA Res.* 23, 1–10. doi: 10.1093/dnares/dsv031
- Zimin, A. V., and Salzberg, S. L. (2020). The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput. Biol.* 16:e1007981. doi: 10.1371/journal.pcbi.1007981
- Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L., and Yorke, J. A. (2013). The MaSuRCA genome assembler. *Bioinformatics* 29, 2669–2677. doi: 10.1093/bioinformatics/btt476