Repeated loss of floral scent in independent transitions from bee to hummingbird pollination in spiral gingers

Kathy Darragh[1,2,+], Kathleen M. Kay[3], Santiago R. Ramírez[1]

[1]Department of Evolution and Ecology, University of California, Davis, CA, USA

[2]Department of Biology, Indiana University, Bloomington, IN, USA

[3]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA

[+]Corresponding author (Email: kdarrag@iu.edu)

ORCID Kathy Darragh 0000-0002-9215-014X

ORCID Kathleen M. Kay 0000-0001-8858-110X

ORCID Santiago R. Ramírez 0000-0003-1306-1315

Keywords

Convergent evolution, floral scent, hummingbird, orchid bee, pollination, spiral gingers, transcriptomics, trait evolution

Abstract

The repeated evolution of the same trait in different lineages provides powerful natural experiments to study the phenotypic and genotypic predictability of how traits are gained and lost. A fascinating example of this is the repeated evolution of hummingbird pollination in plant lineages in the Americas, a widespread and often unidirectional phenomenon. The spiral gingers in the genus *Costus* are ancestrally bee-pollinated and hummingbird pollination has evolved multiple times independently in the tropical Americas. These pollinator transitions are accompanied by predictable morphological and color changes but the changes in floral scent have not been described. In this study, we describe the floral scent composition of 30 species of *Costus* sampled across the phylogeny to understand how floral scent has evolved across the genus with respect to pollinator transitions. We then combine transcriptomics and genomics in one clade of *Costus* to identify genetic expression differences and gene family evolution associated with pollinator transitions. We show that hummingbird-pollinated species have mostly lost their floral scent, whereas bee-pollinated species exhibit either floral scent maintenance or in some cases, gains of more diverse scent profiles. The remarkable consistency of scent loss in hummingbird-pollinated species highlights the shared strong selection pressures experienced by these lineages. Even species with more recent transitions from bee to hummingbird pollination exhibit scent loss, highlighting the rapid breakdown of scent production following pollinator transitions. We also find the floral scent loss appears to be due to gene downregulation rather than pseudogenization. This research highlights the capacity for rapid changes when selection pressures are strong through downregulation of floral scent genes.

Resumen

La evolución independiente de un rasgo en diferentes linajes ofrece la oportunidad de estudiar la predictibilidad de fenotipos y sus bases genéticas. Un ejemplo fascinante es la evolución repetida de polinización por colibríes en varios linajes de plantas en las Américas, un fenómeno generalizado y típicamente unidireccional. Las especies de caña agria del género *Costus* son polinizadas ancestralmente por abejas y la polinización por colibríes ha evolucionado varias veces de forma independiente en América tropical. Cada transición de polinizador ha sido acompañada de varios cambios morfológicos y colores predecibles, pero los cambios en el aroma floral no han sido descritos. En este estudio describimos la composición del aroma floral de 30 especies de *Costus* para comprender cómo el aroma floral ha evolucionado con respecto a las transiciones de polinizadores. Combinamos transcriptómica y genómica en un clado de *Costus* para identificar diferencias en la expresión genética y la evolución de la familia de genes asociadas con las transiciones de polinizadores. Nuestros resultados muestran que las especies polinizadas por colibríes han perdido en su mayoría su aroma floral, mientras que las especies polinizadas por abejas ya sea mantienen el aroma floral y, en algunos casos, muestran ganancias de perfiles de aroma más diversos. La consistencia en la pérdida de olor en especies polinizadas por colibríes resalta las fuertes presiones de selección natural que experimentan estos linajes. Incluso las especies con transiciones más recientes de polinización de abejas a colibríes exhiben pérdida de olor, lo que destaca la rápida interrupción de la producción de olores después de las transiciones de los polinizadores. También encontramos que la pérdida de aroma floral parece estar controlada por regulación génica negativa y no por la pérdida de genes de producción de aromas (*e.g.* pseudogenización). Esta investigación destaca la rapidez de cambios a través de la

regulación negativa de genes que controlan aromas florales cuando las presiones de selección son fuertes.

Introduction

The repeated evolution of the same traits in different lineages is thought to reflect shared selection pressures. These cases of repeated evolution apply not only to the gain of traits, such as the evolution of antifreeze proteins in fish (Chen et al. 1997) or carnivory in plants (Albert et al. 1992), but also the loss of traits, as seen in the eye regression of cavefish (Sifuentes-Romero et al. 2023), or reduction of attractive traits such as flower size or floral scent in self-fertilizing plants (Tsuchimatsu and Fujii 2022). Repeated evolution provides us with systems where we can study the process of adaptation and ask questions about the predictability of evolution, and the extent of constraint, or flexibility, of the underlying molecular mechanisms (Orr 2005; Losos 2011; Sackton and Clark 2019).

One such example of repeated evolution is pollinator transitions in flowering plants. For example, the evolution of hummingbird pollination has occurred many times across flowering plants, and often involves convergence in a suite of traits. Flowers pollinated by hummingbirds are characterized by traits thought to attract and fit hummingbirds (such as a narrow shape and large amounts of nectar) and in addition, traits thought to deter bees (red coloration, lack of nectar guides, lack of scent) (Schemske and Bradshaw 1999; Castellanos et al. 2004; Bergamo et al. 2016). Interestingly, the evolution of hummingbird pollination has primarily been documented in one direction, from ancestral bee pollination to derived hummingbird pollination (Rosas-Guerrero et al. 2014; Abrahamczyk and Renner 2015; Kay and Grossenbacher 2022). However, there is clade-specific variation in this directionality, with the reasons for pollinator transitions remaining unclear (Rosas-Guerrero et al. 2014; Abrahamczyk and Renner 2015; Kay and Grossenbacher 2022; Barreto et al. 2024).

Hummingbird-pollinated flowers lack floral scent (Knudsen et al. 2004; Wessinger 2024). Genomic studies have shown that, even for birds, hummingbirds have a low number of olfactory receptors (Driver and Balakrishnan 2021; Driver 2022), suggesting at least a reduced investment in this sensory modality. Behavioral experiments have also suggested that hummingbirds rely more on visual than olfactory cues when foraging (Goldsmith and Goldsmith 1982). In fact, floral scent analyses of hummingbird pollinated flowers have shown either complete lack of floral scent or low emission rates of floral scent compounds. (Knudsen et al. 2004; Byers et al. 2014; Amrad et al. 2016).

A fascinating group in which to study the evolution of hummingbird pollination is the genus *Costus* (Costaceae). The genus contains over 80 species of perennial understory herbaceous monocots, the majority of which have diverged rapidly within the last 3 million years in Central and South America (Maas 1972, 1977; Vargas et al. 2020). Since establishing in the American tropics, hummingbird pollination has evolved repeatedly and independently in *Costus* from an ancestral state of bee pollination (Vargas et al. 2020). Specifically, bee-pollinated species are pollinated primarily by orchid bees (Apidae: Euglossini), with females visiting flowers for nectar and pollen, and males visiting for nectar (Kay and Schemske 2003). Hummingbird-pollinated species are often pollinated by hermit hummingbirds (Kay and Schemske 2003). While different species were initially classified as hummingbird or bee pollinated (Maas 1972, 1977), many species do not fit the classic paradigm of long tubular red flowers associated with hummingbird pollination (Grant and Grant 1968). For example, many bee-pollinated flowers have some red coloration (Kay and Schemske 2003), and hummingbird-pollinated flowers are short (Yost and Kay 2009). Despite this variation, a single phenotypic optimum is shared across most of the independent transitions to hummingbird pollination in the clade (Kay and Grossenbacher 2022). Traits including small flower size, the absence of nectar guides, and brightly colored floral bracts are all predictive of hummingbird pollination (Kay and Grossenbacher 2022). The combination of recent genomic resources, a well-resolved phylogeny, and the possibility of collecting samples both in the field and in greenhouses, makes *Costus* a powerful system to study the evolution of convergence on hummingbird pollination (Vargas et al. 2020; Kay and Grossenbacher 2022; Harenčár et al. 2023).

While the morphological and color changes involved in the evolution of hummingbird pollination have been well-studied in *Costus*, the expected scent changes have not been characterized. In this study, we describe floral scent in 30 species of *Costus* sampled across the phylogeny and including eight independent origins of hummingbird pollination. We test for differences in floral scent between bee-pollinated and hummingbird-pollinated species and investigate how floral scent has evolved across the genus. We then focus on one clade of *Costus* with variation in floral scent production and generate floral transcriptomic data. We use this data to identify gene expression differences associated with floral scent production, helping us to better understand the genetic changes underlying floral scent evolution in this group. Finally, we annotated a family of genes important for the production of terpenoid floral scent compounds

(the terpene synthase genes (TPS)) in three previously available *Costus* genomes, and a high-quality genome assembly for *Costus allenii* generated in this study, encompassing two independent transitions to hummingbird pollination. High-quality annotations allow us to analyze the evidence of gene loss or pseudogenization following transition to hummingbird pollination in this gene family that plays an important role in floral scent evolution. Our study, therefore, combines multiple approaches to characterize floral scent changes, gene expression changes, and gene family evolution associated with shifts to hummingbird pollination in *Costus*.

Results

<u>Bee-pollinated species produce more compounds and more diverse floral scents</u>

We collected 100 samples from 30 *Costus* species including species previously described as both bee- and hummingbird-pollinated (Fig 1A, Table S1). Using multivariate techniques, we visualized the floral scent profiles to identify the broad patterns in our dataset. Before filtering our data, we found that samples from hummingbird-pollinated species did not differ from ambient control samples (Fig 1B). In contrast, samples from bee-pollinated species spread across a wider section of chemospace, with many samples visually separated from both ambient and hummingbird-pollinated samples (Fig. 1B). We found that both pollination group, and species (nested within pollination group) explained variation in floral scent (PERMANOVA, pollination group: $F_{2,167}=8.0$, $p<0.001$; species: $F_{28,167}=2.5$, $p<0.001$). Interestingly, bee-pollinated species were significantly different from both hummingbird-pollinated species and ambient control samples (Pairwise PERMANOVA, $p=0.003$, whereas hummingbird-pollinated species did not differ from ambient control samples (Pairwise PERMANOVA, $p=0.07$) (Figure S1).

After filtering likely contaminants from the ambient air, we obtained a final dataset with 60 floral scent compounds (Table S2). The majority of these compounds were terpenoid (87%, 52/60), with a few aromatics and compounds of unknown class. The dataset included both monoterpenes and sesquiterpenes. No one compound class was restricted to a specific pollinator group. After removing the ambient compounds, we found that one bee-pollinated sample and eighteen hummingbird-pollinated samples had no detectable floral scent, confirming the similarity of the hummingbird-pollinated samples to the ambient air samples. The unscented hummingbird-pollinated samples were from six species. For two species, *C. montanus* and *C. stenophyllus*, all individuals produced no detectable floral scent (4/4 individuals). Furthermore,

most C. chartaceus (4/5), *C. lasius* (4/7), and *C. wilsonii* (5/8), and a single *C. productus* individual (1/5) lacked floral scent.

We found that bee-pollinated flowers produced more floral scent compounds, averaging 9 (s.d. 10) compounds in comparison to only 2 (s.d. 4) for hummingbird-pollinated flowers (Figure 1C). Bee-pollinated flowers contained more compounds than hummingbird-pollinated flowers (Fig 2A). Similar to the number of compounds, we used a measure of biochemical and structural diversity of the compounds present in a sample to calculate the overall chemical diversity. The Functional Hill Diversity index considers not only the total number of compounds but also their biosynthetic complexity. We found that the bee-pollinated floral scents had a higher chemical diversity (Fig 2B). Bee-pollinated flowers also produced a higher abundance of floral scent, as measured by total ion abundance in each sample, a proxy for compound amount (Fig. 2C).

To see how individual compounds varied between bee- and hummingbird-pollinated species we calculated the percentage of individuals in each pollination group that produced the most common floral scent compounds in the dataset. We found that these compounds are present in a higher proportion of bee-pollinated flowers than hummingbird-pollinated flowers (Fig. 2D). In particular, (*E*)-beta-ocimene and beta-caryophyllene show the largest percentage differences between the two pollination groups (Fig. 2D) and are the only two compounds to exhibit a significant difference in frequency between pollinator groups (Table S2).

To facilitate visualization of our final filtered dataset using NMDS we removed the nineteen samples with no detectable floral scent. When this filtered dataset was plotted, we observed more overlap between the hummingbird- and bee-pollinated samples, but the two sets of samples significantly different (PERMANOVA, pollination group: $F_{1,80}=4.2$, $p<0.001$; species: $F_{26,80}=2.4$, $p<0.001$; Pairwise PERMANOVA bee- vs hummingbird-pollinated species $p=0.001$) (Figure S1).

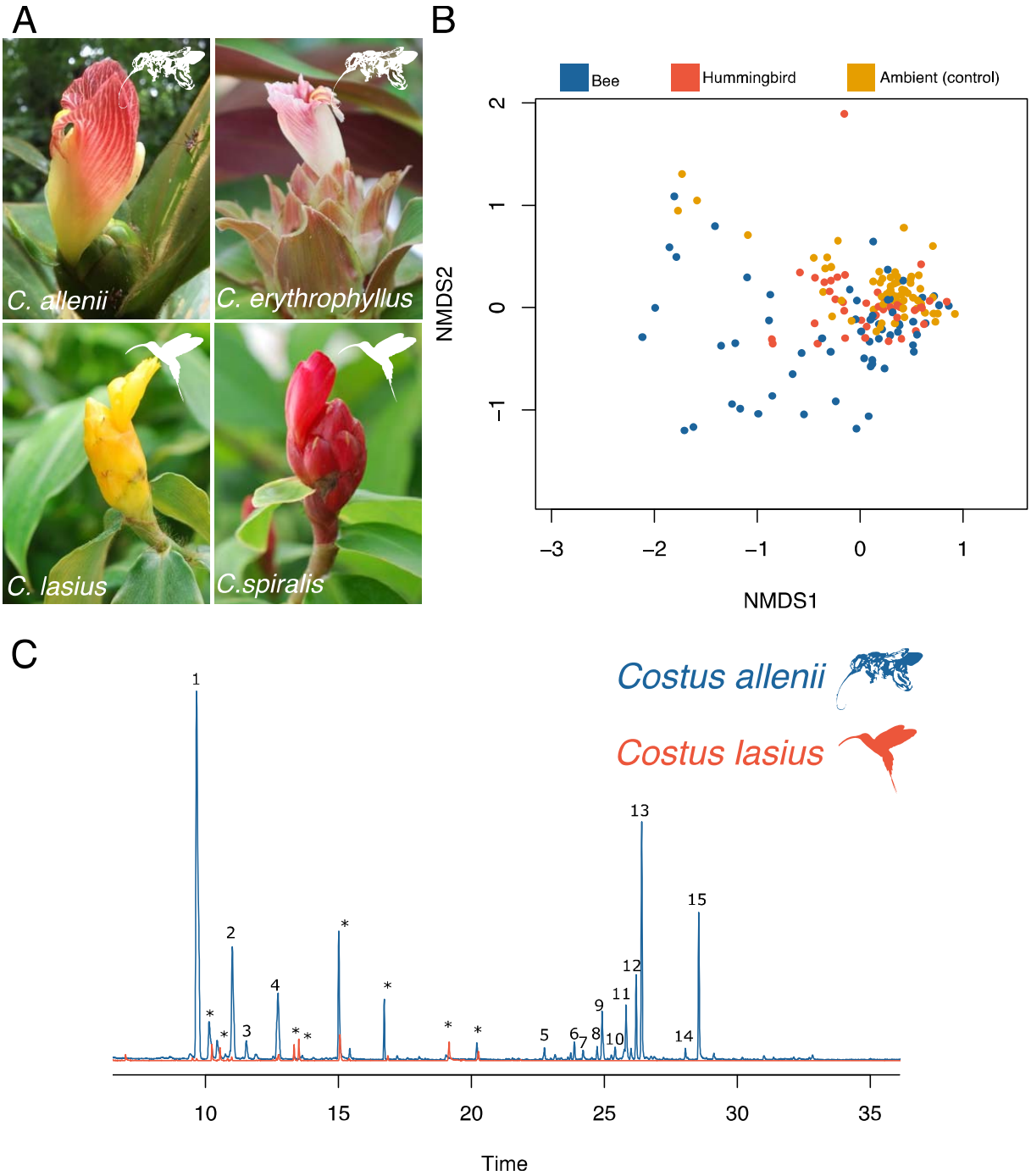

Figure 1. A) Representative images of two pairs of closely related bee- and hummingbird-pollinated species: *Costus allenii* and *C. lasius* and *C. erythrophyllus* and *C. spiralis*. B) NMDS (nonmetric multidimensional scaling) plot illustrating in the variation in the unfiltered dataset (Stress=0.12). Points are labelled as bee-pollinated, hummingbird-pollinated, or ambient samples. Ambient samples are control samples of the air taken in the

same area as the focal plant individual. C) Total ion chromatogram of floral scent from *Costus allenii* and *C. lasius*. We chose the individuals with the second-highest amount of scent from each species. The *C. lasius* sample was shifted on the x-axis by 1.5 minutes to match the *C. allenii* sample due to changes in the GC/MS column over time. Abundance is scaled to the highest peak present in *C. allenii*. 1, alpha-pinene; 2, beta-pinene; 3, beta-myrcene; 4, cineole; 5, alpha-copaene; 6, (*E*)-beta-caryophyllene; 7, gamma-elemene; 8, alpha-humulene; 9, (*Z*)-beta-caryophyllene; 10, germacrene D; 11, alpha-elemene; 12, gamma-cadinene; 13, delta-cadinene; 14, germacrene-d-4-ol; 15, unidentified sesquiterpene (RI=1616). Contaminants are identified with an asterisk (*).

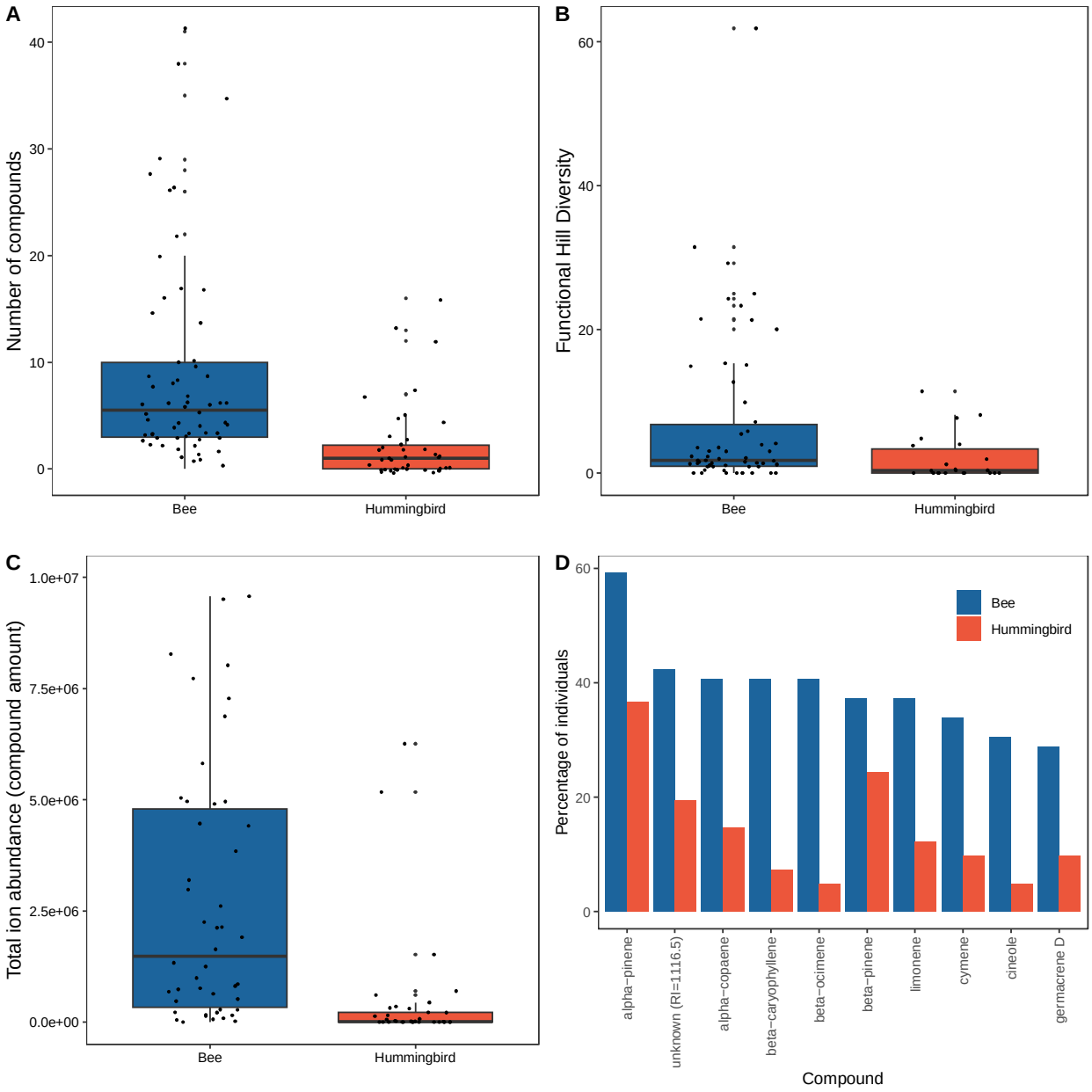

Figure 2. Differences in bee-pollinated and hummingbird-pollinated *Costus* species in terms of A) Number of scent compounds produced (ANOVA, pollination group: $F_{1,70}$=60.4, p<0.0001, species: $F_{28,70}$=9.4, p<0.0001), B) The diversity of scent compounds produced (ANOVA, pollination group: $F_{1,51}$=5.3, p=0.03, species: $F_{27,51}$=1.4, p=0.1), C) Total compound amount produced (measured by proxy of total ion abundance) (ANOVA, pollination group: $F_{1,70}$=109, p=<0.0001, species: $F_{28,70}$=2.8, p<0.001) and D) the percentage of individuals that produced the ten most common floral scent compounds in the dataset.

Significant differences in frequency between bee- and hummingbird-pollinated flowers as determined by Chi-squared testing are highlighted with an asterisk. Fifteen individuals with the highest compound production were removed from C for visualization.

Ancestral state reconstructions show that hummingbird-pollinated species have lost floral scent

To understand the evolution of floral scent in *Costus*, we carried out phylogenetic analyses using a previously published phylogeny (Vargas et al. 2020). In our dataset, we found a strong correlation between pollination group and the phylogenetic covariance matrix using two-block partial least squares analysis (rPLS = 0.56; P=0.01). This reduces the statistical power of testing group differences using phylogenetic ANOVA due to the high aggregation of pollinator groups on the phylogeny (Adams and Collyer 2018). We tested for evidence of phylogenetic signal in compound number, compound diversity, compound amount (total area of floral scent peaks), (*E*)-beta-ocimene production, and (*E*)-beta-caryophyllene production, detecting weak and non-significant phylogenetic signal overall (λ= 4.14795e-05, p=1). We also tested for evidence of phylogenetic signal in the multivariate floral scent dataset, detecting non-significant almost zero phylogenetic signal (Kmult=0.43, p=0.68). We therefore decided against using phylogenetic ANOVA with simulations based on Brownian motion and instead used a method of randomizing residuals in a permutation procedure (Adams and Collyer 2018). We found that compound number, compound diversity, compound amount, and (*E*)-beta-ocimene production were all significantly different between bee-pollinated and hummingbird-pollinated species (compound number, $F_{1,26}$=7.2, p=0.008; compound diversity, $F_{1,26}$=5.8, p=0.01; compound amount, $F_{1,26}$=627.5, p=0.008; (*E*)-beta-ocimene production, $F_{1,26}$=4.7, p=0.03). (*E*)-beta-caryophyllene production did not differ between pollinator groups ($F_{1,26}$=1.0, p=0.4). Our comparative analyses complement our previous analyses, again showing that bee-pollinated species produce higher amounts of more diverse and complex floral scents and have higher (*E*)-beta-ocimene production.

To further investigate how floral scent has evolved across the *Costus* phylogeny, we carried out maximum likelihood ancestral state reconstruction of mean compound production. We overlaid the eight independent origins of hummingbird pollination represented by our dataset as determined in (Vargas et al. 2020). The ancestral state of all *Costus* species found in tropical America is to produce more compounds than average hummingbird-pollinated species (4.5 vs 2).

In this reconstruction, the ancestral *Costus* population produced an average floral scent – more compounds than hummingbird-pollinated species but less than the average bee-pollinated species (Figure 3). Many hummingbird-pollinated species have since completely lost floral scent, as shown by the white circles (Figure 3). We see evidence for scent reduction or loss in all eight origins of hummingbird pollination in our dataset. In contrast, some bee-pollinated species, such as *Costus allenii*, have evolved much stronger floral scents than the ancestral state.

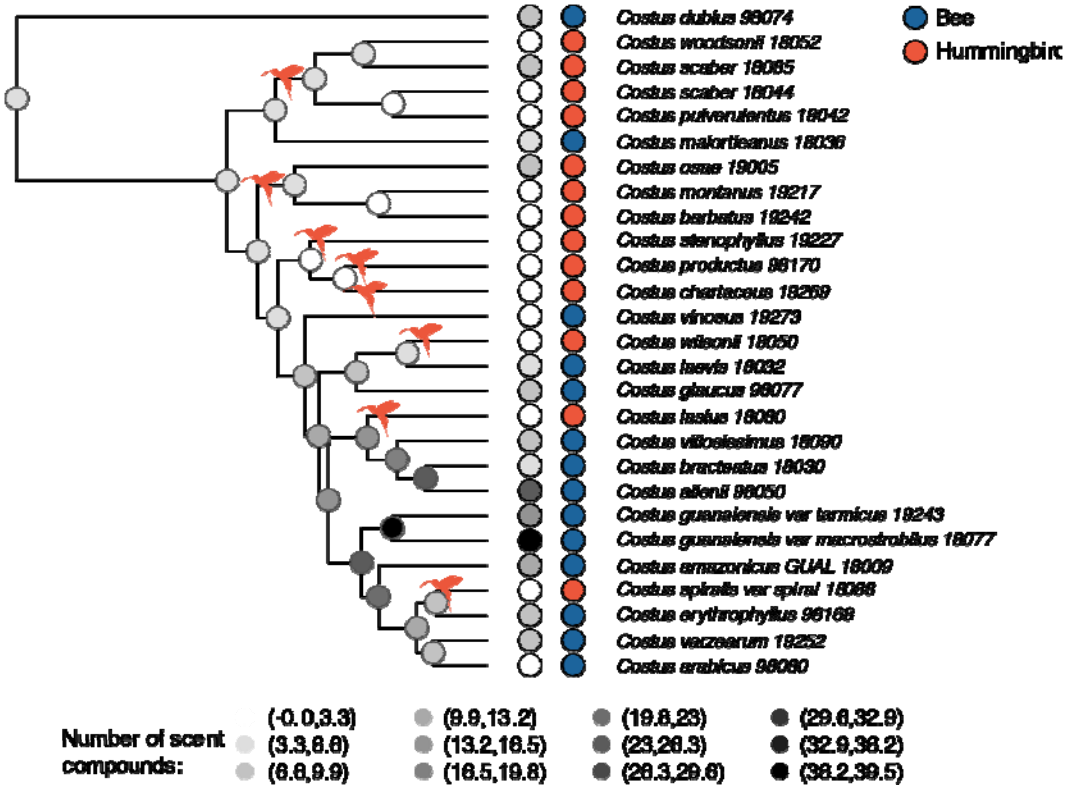

Figure 3. Ancestral state reconstruction of floral scent in the genus *Costus*. Pollinator for each species is illustrated by the right-hand dots. Mean compound number produced by each species is shown by the left-hand dots with each shade of gray representing a range of means. Darker dots indicate more compounds are produced. Predicted compound production at each node is shown following the same color scale. Origins of hummingbird pollination (as noted by hummingbird symbols on branches) and tip labels following (Vargas et al. 2020).

<u>Loss of floral scent is correlated with downregulation of floral terpene synthase genes.</u>

To analyze the genetic basis of floral scent evolution, we first chose a clade of closely related species with variation in floral scent production and pollination systems. We chose the

clade containing *Costus lasius* due to the variation in floral scent production and the availability of a high-quality genome. In this clade, *Costus lasius* (hummingbird-pollinated) produces on average three floral scent compounds, whereas *Costus allenii* (bee-pollinated) produces 24 (one of the highest in our dataset), and *Costus villosissimus* (bee-pollinated) produces seven compounds (Figure 4A).

We first produced floral transcriptomes of all three species. After clustering transcripts with over 95% similarity, we found 112998 transcripts present in *C. allenii*, 80208 in *C. lasius*, and 96886 in *C. villosissimus*. Re-mapping rates were high in all three species: 99.2% in *C. allenii*, 99% in *C. lasius*, and 99.1% in *C. villosissimus*. All three also had BUSCO (Benchmarking Universal Single-Copy Orthologs) scores of more than 80% (Table S3), suggesting high-quality transcriptomes with most single-copy orthologs from the Embryophyta database present in all species.

Due to the dominance of terpenoid compounds in the floral scent, we focused on the expression of terpene synthase genes (TPS). TPS genes are responsible for producing terpenoid compounds. We found that 20 TPS genes were expressed in the floral transcriptome of *C. allenii*, compared to five in *C. villosissimus*, and only one in *C. lasius*, when expression was defined as one transcript per million (TPM) in at least one sample of the species. Furthermore, we found that counts of TPS expression per TPS gene expressed were much higher in both *C. allenii* and *C. villosissimus* when compared to *C. lasius* (Fig. 4B). In fact, in *C. allenii*, a TPS is the highest expressed gene in the floral transcriptome, and TPSs are also the 4th and 6th most expressed genes in the transcriptome, highlighting the extent of expression. On average, 12% of total expression can be attributed to TPSs in *C. allenii*, 0.17% in *C. villosissimus* and 0.00014% in *C. lasius*, highlighting the orders of magnitude differences between species.

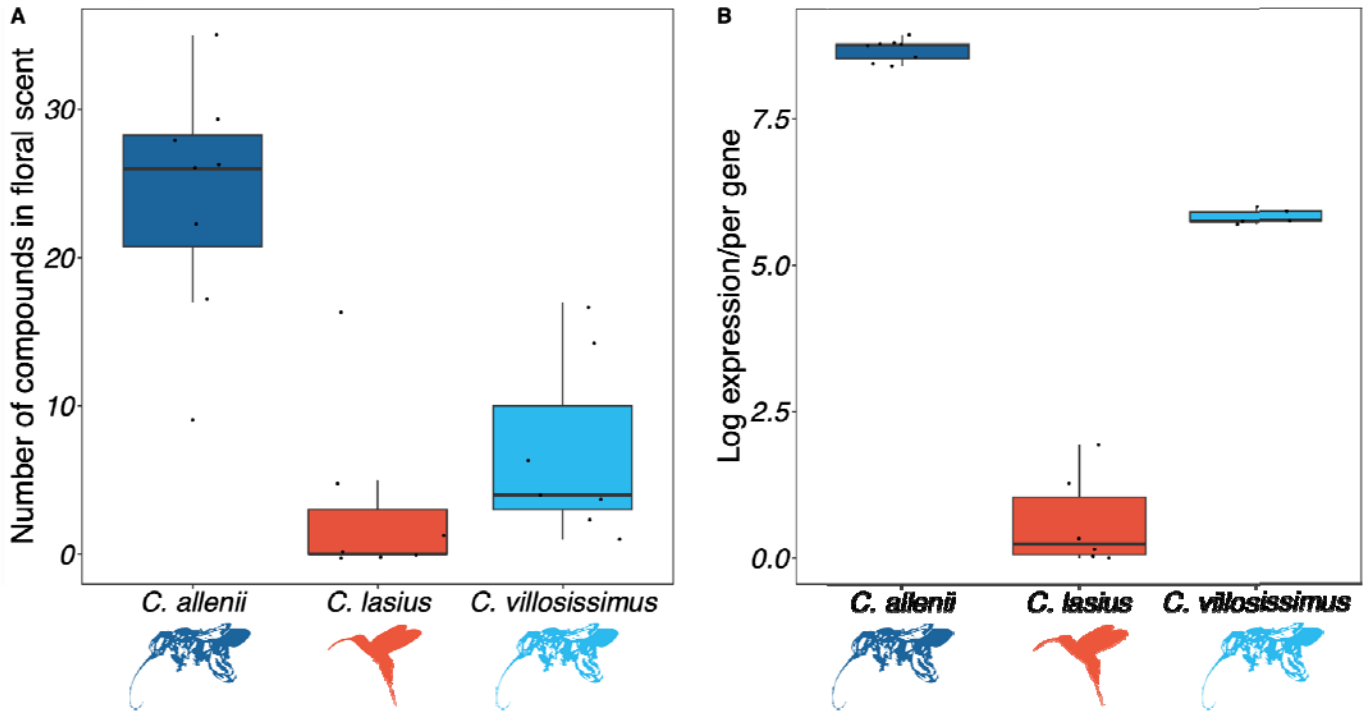

Figure 4. Differences in floral scent and gene expression between *C. allenii*, *C. lasius*, and *C. villosissimus*. Species differ significantly in A) number of floral scent components (ANOVA, $F_{2,19}$=20.9, p<0.001), and (B) expression of terpene synthase transcripts (ANOVA, $F_{2,16}$=508.4, p<0.001).

Floral scent loss is not driven by gene loss in *Costus*

While the transcriptome assemblies can give us insight into levels of TPS expression, it is hard to determine gene copy number due to the redundancies in the assemblies. For this, we need genome data to allow us to determine the number of TPS genes in each species. There are currently three publicly available *Costus* genomes: *C. bracteatus* (bee), *C. lasius (hummingbird)*, and *C. pulverulentus* (hummingbird) (Valderrama et al. 2022; Harenčár et al. 2023). *Costus lasius* and *C. pulverulentus* represent two independent transitions to hummingbird pollination (Kay and Grossenbacher 2022). To generate a dataset of two hummingbird- and two bee-pollinated species, we assembled an additional high-quality reference genome for bee-pollinated *C. allenii* (Table S1 for assembly statistics, Figure S2 to see alignment with *C. lasius* genome). We generated high-quality whole genome annotations for each species with high completeness and consistency (Table 2).

We annotated between 63-75% of each genome as repeat elements (Tables S4, S5, S6, S7). We found that, of the elements that could be classified, retrotransposable elements represented the largest percentage of the genome. These are the most abundant form of repeat elements found in angiosperm genomes. Many repeats were not classified into previously known families, perhaps suggesting the existence of new repeat elements.

Table 1. *Costus allenii* genome assembly statistics

N50 (Mb)	120
L50	5
Total length (Gbp)	1.137
GC%	40.07
Number of Ns per 100 kbp	0
Longest contig/scaffold (Mbp)	143
Number of contigs/scaffolds	423
BUSCO complete	98.8% (1594/1614)
BUSCO single copy	90.8% (1465/1614)

Table 2. BUSCO scores for high quality annotations of each *Costus* genome searched against the Embryophyta database.

	C. allenii	*C. bracteatus*	*C. lasius*	*C. pulverulentus*
Complete	1493 (92.5%)	1457 (90.2%)	1480 (91.7%)	1466 (90.9%)
Complete and single copy	1139 (70.6%)	1116 (69.1%)	1121 (69.5%)	1152 (71.4%%)
Complete and duplicated	354 (21.9%)	341 (21.1%)	359 (22.2%)	314 (19.5%)
Fragmented	12 (0.7%)	25 (1.5%)	19 (1.2%)	21 (1.3%)
Missing	109 (6.8%)	132 (8.3%)	115 (7.1%)	127 (7.8%)
Total searched	1614	1614	1614	1614
Consistent lineage placement (OMArk)	86%	87%	86%	87%
Total gene number	28,442	27,794	28,409	26,696

We annotated a total of 66 TPSs in *C. allenii*, 44 in *C. bracteatus*, 40 in *C. lasius*, and 44 in *C. pulverulentus*. All TPSs belonged to previously described TPS families TPS-a, TPS-b, TPS-c, TPS-e/f, and TPS-g (Fig. 5). Those TPSs responsible for floral scent biosynthesis are primarily found in the TPS-a, TPS-b, and TPS-g families. TPS-a members tend to produce monoterpenes, TPS-b members sesquiterpenes, and TPS-g members either monoterpenes or sesquiterpenes. We identified lineage-specific duplications in both TPS-a and TPS-b, as well as TPS-e/f in the bee-pollinated *C. allenii*. We did not find any evidence for pseudogenization or gene loss in the two hummingbird-pollinated species included in the gene family analyses, *C. lasius* and *C. pulverulentus*, with TPS genes containing complete protein domains and start codons.

We mapped the RNAseq reads both to the *C. allenii* genome and the *C. lasius* genome. In *C. allenii*, we found a consistent percentage of successfully mapped reads attributed to TPSs as with the transcriptome (12% transcriptome, 13% mapped to either genome). We found some differences for both *C. lasius* and *C. villosissimus*. For the *C. villosissimus* transcriptome mapping, we found 0.17% of reads were attributed to TPSs, in comparison to 0.009% when mapped to either genome. Due to a lack of genome for *C. villosissimus*, it is unclear whether this could be due to genetic changes or rearrangements not captured by the available genomes. For *C. lasius*, we found only 0.00014% of reads mapped to the transcriptome were due to TPSs, in contrast to 0.011% of the genome-mapped reads.

When *C. allenii* reads were mapped to the *C. allenii* genome, we found high levels of multi-mapping. This is due to recent TPS duplications in *C. allenii* which are highly expressed. To illustrate this, the highest gene expressed in *C. allenii* is *TPS49*, and the third most expressed "gene" consists of reads that multimap to *TPS49* and *TPS52*. This clade is represented by only one gene in each of the other species. Other recent lineage-specific duplications include *TPS2*, *TPS3*, *TPS4*, and *TPS5* where the majority of reads multimap due to sequence similarity between duplicates. Both of these clades have only one member in each of the other species with an available genome sequence (Fig S3). This highlights the complexity of mapping to gene families with recent duplications, especially when they are highly expressed, meaning that the typical choice of ignoring multi-mapped reads will lose large percentages of mapped reads from these duplications. We, therefore, decided to carry out our differential expression analysis using the *C. lasius* genome.

By mapping all reads to the *C. lasius* genome, we identified 4790 genes differentially expressed between *C. allenii* and *C. lasius*, 7678 between *C. allenii* and *C. villosissimus*, and 6191 between *C. lasius* and *C. villosissimus*. Of the 40 TPS genes found in the *C. lasius* genome, 15 were differentially expressed between *C. allenii* and *C. lasius*, with 13 of these upregulated in *C. allenii*. This includes *TPS33*, orthologous to the *TPS49*/*52* clade, the most highly expressed genes in *C. allenii*. We found 19 TPSs that were differentially expressed between *C. allenii* and *C. villosissimus*, with 14 upregulated in *C. allenii* and 5 in *C. villosissimus*. Ten TPS genes were differentially expressed between *C. lasius* and C. *villosissimus*, with 7 upregulated in *C. villosissimus* and 3 in *C. lasius*. Overall, more TPSs were upregulated in *C. allenii* when compared to either *C. lasius* or *C. villosissimus*. Our gene ontology analysis found little evidence for functional enrichment in differentially expressed genes between species (Table S8).

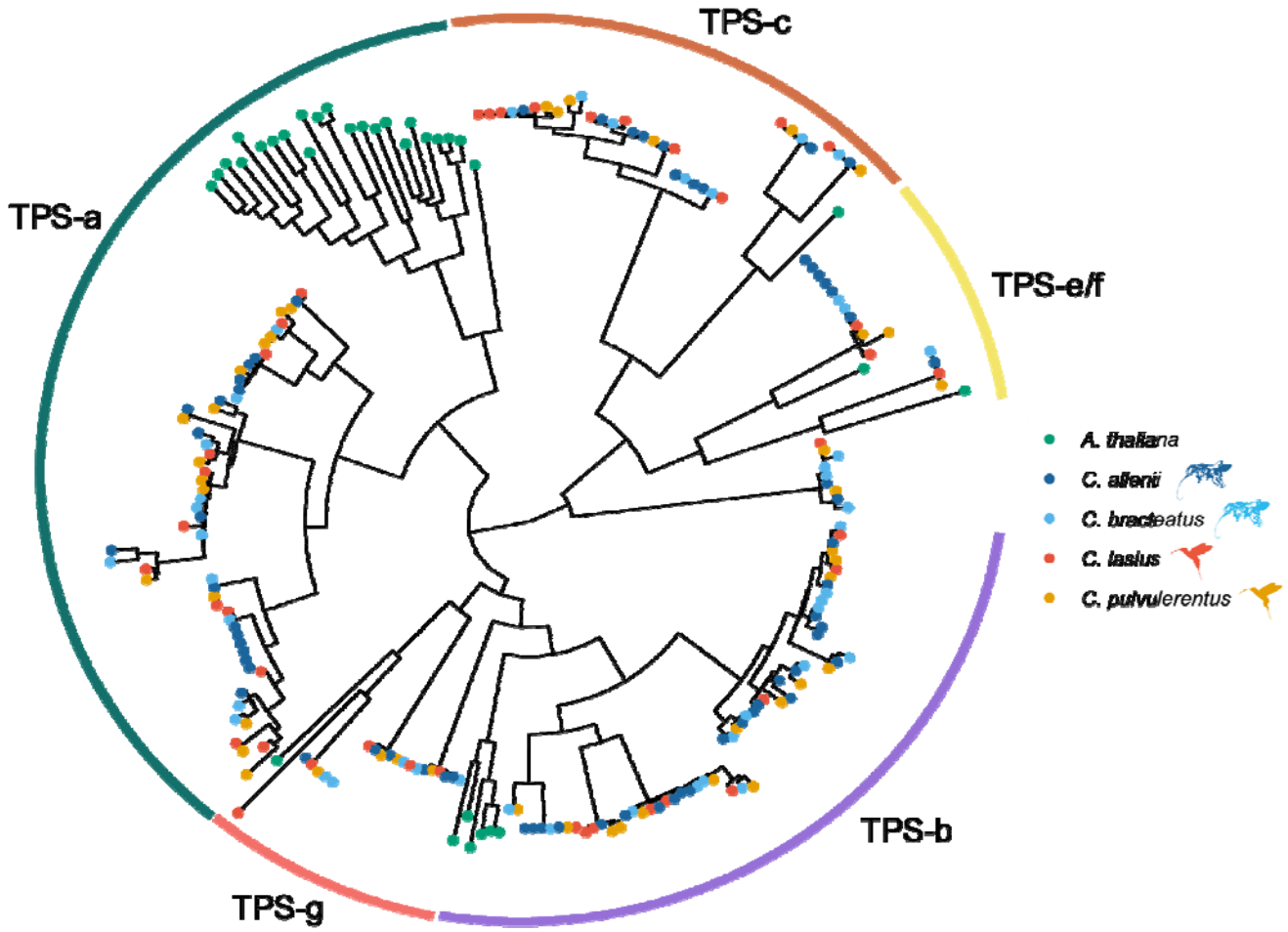

Figure 5. Phylogenetic analysis of terpene synthase gene families. The phylogeny was constructed in IQ-TREE with 1000 bootstraps and model LG. TPS family classifications based on *Arabidopsis thaliana* are illustrated.

Discussion

Repeated evolutionary transitions provide unique opportunities to study the extent of convergence both at the phenotypic and molecular level, shedding light on the predictability of evolution. Transitions to hummingbird pollination are widespread in the Americas and are predicted to be associated with loss of floral scent, however, this has not been studied in a comparative framework across multiple independent transitions within the same genus. In this

study, we combine comparative chemical ecology with genomics and transcriptomics and demonstrate that floral scent evolution in the genus *Costus* is highly associated with pollinator type. We found that hummingbird-pollinated *Costus* species have lost or reduced their floral scent across eight independent evolutionary origins, whereas bee-pollinated species in the Americas exhibit either maintenance of floral scent or gain of a more diverse scent profile. We find that this is related to downregulation of genes associated with floral scent production (in this case terpene synthases) in floral tissues of hummingbird-pollinated species but is not accompanied by loss of terpene synthase genes.

Bee-pollinated *Costus* species showed much greater variation in their floral scents than hummingbird-pollinated species, with some having extremely diverse floral scents, yet others producing far fewer compounds. Bee-pollinated species also show greater morphological variation, possibly reflecting variation in orchid bee diversity (Kay and Grossenbacher 2022). Furthermore, not all species in our dataset follow the expected patterns. For example, *Costus arabicus* and *Costus vinosus* (bee-pollinated) produce little scent. It is unclear whether the variation in floral scent between different bee-pollinated species is related to pollinator preference, other abiotic or biotic factors, or historical contingency.

There are multiple hypotheses for why floral scent would be lost in hummingbird-pollinated species. Firstly, it could be because there is simply no selection to maintain the production due to poor sense of smell, or lack of preference shown for floral scent, by hummingbirds (Goldsmith and Goldsmith 1982; Knudsen et al. 2004; Driver and Balakrishnan 2021). If this were the case, we might expect to find a correlation between length of time following pollinator transition and extent of floral scent loss. However, our data shows that downregulation of floral scent can occur on relatively quick evolutionary timescales. *Costus wilsonii*, *C. lasius*, and *C. spiralis* are all hummingbird-pollinated species that are most closely related to bee-pollinated species, representing more recent evolutionary shifts to hummingbird pollination. Yet all three species produce little to no floral scent, suggesting there can be quick breakdown of floral scent production following pollinator transitions. This observation is consistent with previous studies that have shown that floral scent often evolves rapidly in response to selection (Ramos and Schiestl 2019; Zu et al. 2020; Liu et al. 2024).

The rapid and repeated loss of scent in hummingbird-pollinated species suggests a shared strong selection pressure against floral scent production. One potential mechanism is the energetic cost of terpene production. A similar argument has been made in the case of eye regression in cavefish. If vision is energetically costly to maintain, then it should be selected against when no longer useful (Moran et al. 2015; Sifuentes-Romero et al. 2023). This is supported by evidence of the energetic costs of maintaining a visual system (Moran et al. 2015). The metabolic cost of terpene production is unknown, but is thought to be minimal (Raguso 2016; Pichersky and Raguso 2018). In fact, it has been argued that the complex network of terpene metabolism, in which individual enzymes produce multiple products, makes terpene pathways more efficient, compared with a linear pathway with separate enzymes and regulators for each step (Lanier et al. 2023).

An alternative scenario to the energetic cost is that terpene production has ecological costs in hummingbird-pollinated species. For example, there may be strong selection against bee attraction (Castellanos et al. 2004; Bergamo et al. 2016). Many benefits to hummingbird, compared with bee, pollination have been described. Birds do not consume pollen leading to less pollen loss, and birds can travel longer distances between plants and promote higher gene flow, which is beneficial to combat inbreeding depression (Thomson and Wilson 2008; Krauss et al. 2017; Dellinger et al. 2022). One caveat is that comparisons in pollination effectiveness between bees and hummingbirds have not been carried out in orchid bees. Both male and female orchid bees tend to forage over long distances, perhaps resulting in a more “bird-like” pollen transport (Janzen 1971, 1981; Wikelski et al. 2010; Opedal et al. 2017; Kay and Grossenbacher 2022; Gamba and Muchhala 2023). While pollinator efficiency is unknown in *Costus*, rapid floral scent loss is congruent with the hypothesis of “bee-avoidance” following establishment of hummingbird pollination.

A final piece of evidence to support bee-repellence is the large reduction in beta-ocimene production in hummingbird-pollinated flowers relative to bee-pollinated flowers, as this compound is a common insect pollinator attractant (Farré-Armengol et al. 2017). In *Mimulus*, loss of monoterpenes, including beta-ocimene, maintains reproductive isolation between bee- and hummingbird-pollinated species (Byers et al. 2014). Interestingly, we found that α- and β-pinene were produced by both bee- and hummingbird-pollinated species. One potential reason is that pinenes do not repel orchid bees but do repel other types of bees. Terpenoid-dominated

scents have been found to repel honey bees (Larue et al. 2016), and in particular α-pinene has a repellent effect on honey bees (Fernandes et al. 2019). We speculate that bee-pollinated species emit a mix of compounds, some of which potentially attract orchid bees, and some of which potentially repel other pollinators to narrow attraction to orchid bees, specifically. Then when floral scent is reduced in hummingbird-pollinated flowers, the compounds which are attractive to orchid bees are lost but not those which are general bee repellants. Further experimental work is needed to test these hypotheses.

We found that the floral scents of *Costus* are dominated by terpene compounds. Terpene diversity in a given organism is mainly determined by the terpene synthase (TPS) gene family, which carry out crucial steps in terpene formation (Tholl 2006). Neither hummingbird-pollinated species with an available genome sequence (*C. lasius* and C. *pulverulentus*) exhibited loss of TPS genes. There are multiple reasons why trait loss would be associated with changes in gene regulation rather than gene loss. Firstly, these cases may represent recent transitions to hummingbird pollination, and at least some of these TPSs that were responsible for floral scent production will eventually be lost. This has been proposed as a "step-wise" model to pseudogenization, whereby reduction in expression is followed by relaxed selection and then loss-of-function mutations, as seen in an enzyme lost in marine mammals (Graham et al. 2023).

Alternatively, we might not expect any TPS genes to be lost even in the long term. Terpenes, and therefore TPSs, are important for many other roles in plants. Plants use terpenes to deter herbivores, to attract parasitoids or predators of herbivores, and as anti-fungal and anti-bacterial compounds to prevent disease (Gershenzon and Dudareva 2007). These pleiotropic functions of TPSs could explain why we observe downregulation in floral tissues specifically rather than any pseudogenization. One argument against this is that in other systems TPS pseudogenization does occur following pollinator transitions. In *Mimulus*, loss of monoterpene production (specifically beta-ocimene), is not due to expression differences but loss-of-function mutations (Peng et al. 2017). It is unclear whether pseudogenization will occur at some point in *Costus*.

Although we did not observe gene loss in hummingbird-pollinated lineages, we did observe extensive lineage-specific duplications in the bee-pollinated species *C. allenii*. *Costus allenii* was one of the species with highest scent diversity in our dataset, and we found multiple

TPS genes to exhibit several recent lineage-specific duplication events. We also found extremely high TPS expression, accounting for approximately 12% of expression in floral tissue. The gene with the highest expression in floral tissue was found to be a TPS-a family member, with high expression also found also for several TPS-b members. These two families produce sesquiterpenes and monoterpenes, respectively, both of which are found in the floral scent of *C. allenii*. In *C. lasius* and *C. villosissimus*, we found reduced TPS expression compared to *C. allenii*, as expected based off their reduced floral scent production. We also saw high levels of phenotypic variation in *C. villosissimus*, and there is no publicly available genome, making interpretation of the results more difficult. We speculate that *C. allenii* has higher floral scent production than closely related *C. villosissimus* because *C. allenii* grows as isolated individuals in deeply shaded, perennially wet forest, whereas *C. villosissimus* grows at forest edges and large treefall gaps, and often has conspecifics nearby (Chen and Schemske 2015). Thus *C. allenii* may need more scent to attract pollinators. While the overall trends are clear, we hesitate to over-interpret the gene expression data as we have only sampled at one timepoint and it is possible that there is floral scent production happening earlier that is not captured in our data.

We observe more consistency and predictability in the phenotypic patterns than the underlying molecular mechanisms. In many systems, parallel changes in floral scent have accompanied independent pollinator shifts, including the current study (Knudsen et al. 2004; Byers et al. 2014). This can also be seen in other reproductive shifts, such as transitions to self-fertilization (Wozniak et al. 2022). In some systems, the underlying molecular mechanisms of floral scent loss are also convergent, having been shown to involve the same loci in multiple independent evolutionary events. For example, loss of floral scent in different *Mimulus* lineages is due to changes at the same loci, including frameshifts, deletions, and potentially post-transcriptional regulation (Peng et al. 2017). A similar pattern was found in two different systems that have lost benzaldehyde floral scent production due to either evolution of self-fertilization (Sas et al. 2016), or transition to hummingbird pollination (Amrad et al. 2016). The same locus contains loss-of-function mutations in both cases (Raguso 2016). However, this is not universally true. In *Capsella*, the loss of beta-ocimene following transition to self-fertilization in one lineage is due to changes in subcellular enzyme localization, but QTL results suggests another locus is responsible for scent loss in a related species (Wozniak et al. 2022). In this study, we saw no evidence for pseudogenization. Instead, we found evidence that gene expression differences are

responsible for scent differences between *Costus* species, with scent loss likely due to changes in, potentially tissue-specific, gene expression.

In summary, we have shown that *Costus* species exhibit repeated loss of floral scent following transitions from bee to hummingbird pollination, suggesting a shared strong selection pressure favoring scent loss. We also find bee species that have recently gained incredibly diverse scents, which, at least in *C. allenii*, is associated with lineage-specific terpene synthase duplications. We show that scent loss is not associated with loss of terpene synthase genes but rather appears to be due to gene downregulation. This demonstrates the capacity for rapid metabolic changes in response to selection following pollinator transitions. This system provides us with the exciting opportunity to identify the cis- or trans-regulatory changes driving gene expression differences in multiple independent transitions to hummingbird pollination, allowing us to study the extent of genetic convergence.

Materials and Methods

Floral scent sample collection

We collected 100 samples of 30 species of *Costus* between 2019 and 2023 (Table S1). All samples for *Costus glaucus*, *Costus laevis*, *Costus montanus*, *Costus stenophyllus*, and *Costus wilsonii* were collected in Costa Rica (La Gamba, Las Cruces, Las Alturas, and Monteverde). All other species were sampled in the UCSC greenhouses in Santa Cruz, California, USA from plants composing the UCSC living *Costus* collection. For each day of sampling, both floral scent samples and ambient samples were collected (with one exception: a *Costus scaber* sample collected on April 8th 2019 was missing a control and so we used a control from the nearest date available, March 7th). Scent was sampled using dynamic headspace sampling from approximately 9 am until 3 pm. Flowers from living plants were placed inside oven bags (Reynold's). A Spectrex PAS-500 air pump with Tygon tubing was used to pull scented-air through a volatile trap for 6 hours at a flow rate of 200 ml air / min. The volatile trap contained a filter of glass wool and 50 mg Porapak. We conditioned scent traps by passing 5 mL of hexane before usage. Volatiles were eluted from the trap using 400µl of hexane and stored at -20°C.

Chemical analysis

Samples were analyzed using Agilent model 5977A mass-selective detector with an Agilent GC model 7890B. Samples were run on a HP-5 Ultra Inert column (Agilent, 30m x 0.25mm, 0.25 µm). For analysis, one microliter of each sample was injected in splitless mode using the ALS 7694 autosampler with helium as the carrier gas (250°C injector temperature). The program started at 40 °C for 3 minutes and then rose at 5 °C/min to 210°C. The temperature was held at 210°C for 1 minute, before rising at 20°C/min to 300°C where it was held for 2.5 minutes, and 315°C for a final minute.

Compounds were identified by comparing mass spectra and retention indices with reference libraries. Compounds that were not found in at least two samples were removed. A two-step process was used to filter datasets for floral scent components. First, compounds were filtered to include only those which are found in five times higher amounts in floral than ambient samples using the "filter_ambient_ratio" function from the *bouquet* package (Powers et al. 2023). This resulted in a list of ambient compounds which were not considered floral scent compounds. Second, using the raw data matrices, we subtracted the ambient compound amounts from the corresponding floral scent amount collected on the same date. We then removed the ambient compounds identified by the package *bouquet* from this subtracted matrix to create the final data matrix of floral scent.

Statistical analyses of floral scent

To visualize divergence in floral scent among samples we used nonmetric multidimensional scaling (NMDS) using the "metaMDS" function in *vegan* (Oksanen et al. 2020). To test for differences both between pollination groups, and between floral and ambient samples, we carried out a PERMANOVA (permutational multivariate analysis of variance) using the "adonis2" function in *vegan* (Bray–Curtis distance matrix, 1000 permutations). To identify which groups were significantly different, we carried out post hoc pairwise testing using the "pairwise.perm.MANOVA" function in the *RVAideMemoire* package with a Bonferroni correction (Hervé 2021). We did this for both filtered and unfiltered datasets.

We tested for differences in compound number produced by bee-pollinated and hummingbird-pollinated species using ANOVA with a nested model including species within pollinator group. To test for differences in the diversity of compounds produced by bee-pollinated and hummingbird-pollinated species, we used the package *chemodiv* (Petrén et al.

2023). This package considers the biochemical and structural diversity of the compounds present in a sample to calculate the overall chemodiversity per sample. First, we used the "calcDiv" function to calculate the functional Hill diversity index for each sample. Then we carried out an ANOVA test for the diversity indices, using a nested model of species nested within pollinator group to test for differences in chemodiversity. We also tested for differences in compound production using ANOVA. In this case a proxy for compound amount was used: the sum of the total area of floral scent peaks in the sample, or the total ion abundance in the sample. To test for difference in compound frequency between bee- and hummingbird-pollinated samples, we carried out chi-squared tests and adjusted the p-values using the "p.adjust" function in R.

Phylogenetic analyses of floral scent

To investigate the evolution of floral scent in a phylogenetic context we used a previously published phylogeny (Vargas et al. 2020). First, we trimmed the phylogeny to include only those species for which we have scent data using the "drop.tip" function in *ape* (Paradis and Schliep 2018). To test for correlation between pollination group and the phylogenetic covariance matrix, we used a two-block partial least squares analysis using the "two.b.pls" function in the *geomorph* package (Adams and Otárola-Castillo 2013; Adams and Collyer 2018). A significant result means that our power to test for differences between groups using phylogenetic simulation-based is weakened due to group aggregation on the phylogeny. We then used the "phylosig" function in the package *phytools* to test for phylogenetic signal in multiple traits (compound number, compound diversity, compound amount, (*E*)-beta-ocimene production, and (E)-beta-caryophyllene production) (Revell 2012). We also tested if floral scent data overall shows phylogenetic signal using the multivariate generalization of Bloomberg's K (K_{mult}) (Adams 2014). Under Brownian motion, K_{mult} is expected to be 1. If $K_{mult}>1$, higher phylogenetic signal is detected than the null expectation under Brownian motion. If $K_{mult}<1$, lower phylogenetic signal is detected than the null expectation under Brownian motion. To calculate the K_{mult} of *Costus* floral scent, we used the filtered dataset with 1000 iterations to test significance. We carried out the analysis using the "physignal" function in the *geomorph* package (Adams and Otárola-Castillo 2013). Due to the weak phylogenetic signal detected in both multivariate and univariate analyses, we used the "procD.pgls" function in the *RRPP* package to test for differences in traits between pollinator groups rather than a simulation-based ANOVA (Adams and Collyer 2018; Collyer and Adams 2018). This function uses a method of randomizing

residuals in a permutation procedure to evaluate the significance of phenotypic differences between groups considering the phylogeny and has higher statistical power than methods using phylogenetic simulations (Adams and Collyer 2018). We set lambda at 4.14795e-05 as calculated for the univariate variables. We corrected for multiple testing using the "p.adjust" function in R.

To better understand how floral scent has evolved in the genus *Costus*, we carried out an ancestral state reconstruction of the number of compounds produced. We used the function "ace" in the *ape* package for ancestral character estimation (Paradis and Schliep 2018). We divided the compound number trait into twelve categories, each of which were assigned a different shade of blue. We plotted both the compound number for each phylogenetic tip label and internal node to illustrate how floral scent has evolved in *Costus*. We overlaid the origins of hummingbird pollination as described in (Vargas et al. 2020).

Sample collection for RNAseq

Live individuals of *Costus allenii* and *Costus villosissimus* were collected in Soberanía National Park, Colón Province, Panama. Vouchers are deposited at University of Panama (*Costus allenii*: 0132264, 0132265; *Costus villosissimus*: 0132270). Live individuals of *Costus lasius* were collected in El Valle de Antón, Cocle Province, Panama. Voucher is deposited in Michigan State University (MSC) herbarium (Kay 0321). For all three species, the individuals used for RNAseq analyses were grown in greenhouses at UCSC. Freshly opened flowers of each species were removed at approximately 9 am and flash frozen immediately in liquid nitrogen before being stored at -80°C until extraction. Flowers of *Costus* open early in the morning and fall off by mid-afternoon.

RNA extraction and sequencing

Frozen flowers were removed and dissected on dry ice. Tissue from the petals and labella was broken down using a TissueLyser (Qiagen). RNA was then extracted using an RNeasy Mini Kit (Qiagen) with Qiashredder spin columns (Qiagen) following standard protocols. Samples were sent to Novogene for quality control, library preparation, and sequencing. Our final dataset consisted of 19 samples: 8 *Costus allenii* (from 4 individual plants), 6 *Costus lasius* (from 2 individual plants), and 5 *Costus villosissimus* (from 1 individual plant). Samples were sequenced on a NovaSeq 6000 using 150 bp paired end reads. This generated approximately 53 million

reads per library (mean=52.74 million, SD=12.01 million, N=19). The raw sequence data were deposited in the SRA under the BioProject accession PRJNA1010186.

Transcriptome assembly

Before assembling transcriptomes, we trimmed the raw reads using TrimGalore! (Martin 2011). We then assembled transcriptomes for each species using Trinity version 2.5.1 with default settings (Grabherr et al. 2011; Haas et al. 2013). To reduce redundancy, we clustered contigs with a minimum sequence identity of 95% using CD-HIT version 4.6.8 (Fu et al. 2012). We assessed transcriptome completeness and quality using BUSCO (v5.6.1) (embryophyta_odb10 dataset) (Simão et al. 2015; Manni et al. 2021; Zdobnov et al. 2021). We also checked assembly quality by mapping RNAseq reads back to the assembly using Bowtie (v2.4.1) to calculate mapping rates (Langmead and Salzberg 2012).

Terpene synthase gene expression analysis

We extracted the longest isoforms for each Trinity "gene" using the Trinity script "get_longest_isoform_seq_per_trinity_gene.pl" and translated these sequences. We then searched for potential TPSs using hmmsearch two HMM profiles: Terpene_syth_C (PF03936) and Terpene Synthase N-terminal domain (PF01397). We used the R package rhmmer to parse the output from hmmsearch resulting in a list of putative TPS sequences for each species. To reduce the TPSs to only those expressed in the dataset, we used a threshold of TPM>1 (transcripts per million) in at least one individual of a species (Eddy 2011; Arendsee 2017). To calculate expression of each putative TPS we normalized counts using Gene length corrected trimmed mean of M-values (GeTMM), an approach that improves inter- and intra-sample comparisons (Smid et al. 2018). To do this, we first calculated FPKM (fragments per kilobase of transcript per million fragments mapped) for each transcript and carried out TMM-normalization using the package *edgeR* (Robinson et al. 2010). We then combined GeTMM values for transcripts which were identified by Trinity to be the same gene. This does not normalize counts between species but is used to determine whether genes are lowly or highly expressed within a species.

Costus allenii genome assembly

We selected a wild-collected individual *C. allenii* from Pipeline Road, Soberanía National Park, Colón Province, Panama for *de novo* genome assembly. A voucher for this population of *C. allenii* is deposited at the University of Panama Herbarium (PMA:0132264). This individual was grown to flowering in the UCSC greenhouses from a cutting, and we harvested 1.5 g fresh tissue from leaf meristems and flower buds. We isolated cell nuclei following Workman (Workman et al. 2019) and then used the Circulomics Nanobind Plant Nuclei Big DNA kit to extract HMW DNA. We assessed DNA concentration with a Qubit Fluorometer (Thermo Fisher Scientific), purity with 260/280 and 260/230 absorbance ratios from a NanoDrop Spectrophotometer (Thermo Fisher Scientific) and ran a 7% agarose gel at 70 V for over 8 h to assess DNA fragment integrity and size. HMW DNA was sent to the UC Davis Genomics Center for additional cleaning, Pacific Biosciences HiFi library preparation, and sequencing on one PacBio Revio SMRT cell. This produced 74.8 Gb of HiFi yield and 7.9 M HiFi reads. We assembled the PacBio data into a primary and alternative assembly using HiFiasm (v0.19.8)(Cheng et al. 2021). We then used RagTag (v2.1.0)(Alonge et al. 2022) to scaffold *C. allenii* contigs using the closely related *C. lasius* as a reference. We used D-GENIES (Cabanettes and Klopp 2018) with Minimap2 (Li 2018) to generate a dotplot comparing the two assemblies both before and after scaffolding. The genome assembly quality was assessed using BUSCO v 5.6.1 (Simão et al. 2015). Raw sequence data and the final genome assembly were submitted to NCBI under BioProject PRJNA1091680.

Genome annotations

Of the three species for which we generated RNAseq data, an unannotated genome was previously available for *C. lasius* (Harenčár et al. 2023). In this study we also generated a genome assembly for *C.allenii*. In addition, previously published genomes are available for both *C. bracteatus* and the more distantly-related *C. pulverulentus* (Valderrama et al. 2022; Harenčár et al. 2023). We started by processing RNAseq data as evidence for annotations. Trimmed reads from *C. lasius* were mapped using 2pass mapping in STAR to the *C. lasius* genome, and the *C. pulverulentus* genomes (Martin 2011; Dobin et al. 2013). Bam files were concatenated for use as described below. Trimmed reads from *C. allenii* were mapped using 2pass mapping in STAR to both the *C. allenii* and the closely related *C. bracteatus* genomes, and bam files concatenated per species (Martin 2011; Dobin et al. 2013).

We annotated the genomes using BRAKER3 (v3.0.7) which combines RNA-seq and protein data in an automated pipeline (Lomsadze et al. 2005, 2014; Stanke et al. 2006, 2008; Gotoh 2008; Iwata and Gotoh 2012; Buchfink et al. 2015; Hoff et al. 2016, 2019; Kovaka et al. 2019; Brůna et al. 2020, 2021; Pertea and Pertea 2020; Gabriel et al. 2023). This pipeline requires a masked genome. We identified and annotated tandem repeats and transposable elements using RepeatModeler (Flynn et al. 2020) (v.2.0.4 for *C. lasius* and v.2.0.5 for the other genomes) and RepeatMasker v.4.1.5 (Tarailo-Graovac and Chen 2009). We then used RepeatMasker with the outputs of RepeatModeler to generate softmasked genomes. The masked genomes were then provided to BRAKER along with plant proteins downloaded from OrthoDB (odb10_plants) (Zdobnov et al. 2021). For all species, BRAKER was run in ETP mode which combines evidence from both proteins and RNAseq (Gabriel et al. 2021). BUSCO v 5.6.1 (Simão et al. 2015) and OMArk (Nevers et al. 2022) were used to evaluate the completeness and consistency of the sets of gene models. During the project, the *C. lasius* genome was updated (direction of arms on chromosome 1 were altered) and so we re-scaffolded chromosome 1 of *C. allenii* and transferred our annotation using liftoff (Shumate and Salzberg 2021).

Terpene synthase gene family annotation

Following whole-genome annotation, we manually curated the gene models for terpene synthase (TPS) genes in the annotation. To do this, we used bitacora (Vizueta et al. 2020), a pipeline which curates an existing annotation and finds additional gene family members. In this pipeline, BLASTP and hmmer are used to search for gene family members of interest within the existing annotation (Altschul et al. 1990; Eddy 2011). Additional regions of the genome are then searched using TBLASTN and are annotated using GeMoMa (Keilwagen et al. 2019). Hmmer is used to validate the gene models. As input, we used two HMM profiles: Terpene_syth_C (PF03936) and Terpene Synthase N-terminal domain (PF01397) downloaded from Pfam. In addition, we provided previously annotated genes from *Arabidopsis thaliana* and *Oryza sativa* (Chen et al. 2011; Yu et al. 2020; Jia et al. 2022). Gene models from bitacora and the genome-wide annotation from BRAKER were compared and TPS genes were manually curated using IGV (Robinson et al. 2011). We also searched for additional TPSs by carrying out exonerate searches using the curated *Costus* gene models (Slater and Birney 2005). We checked all annotated TPSs for complete protein domains using the NCBI conserved domain search (Marchler-Bauer et al. 2015). We excluded proteins with less than 250 amino acids.

Gene family analysis

To classify these TPSs in previously described TPS families, we constructed a phylogenetic tree based on amino acid sequences with TPSs from *Arabidopsis thaliana*. The *Costus* TPS protein sequences, combined with those from *A. thaliana*, were aligned with MAFFT v7.508 (-maxiterate 1000, using L-INS-I algorithm) (Katoh and Standley 2013). We then trimmed the alignment using trimAl (-gt0.6, sites only included when present in 60% of sequences) (Capella-Gutierrez et al. 2009). We used this alignment to construct a gene tree with IQ-TREE and the ModelFinder function to determine the best-fit model (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017; Hoang et al. 2018). We rooted the tree using the *phytools* package in R ("midpoint.root" function)(Revell 2012; R Core Team 2023). We plotted the resulting gene trees using the following packages in R: *ape* (Paradis and Schliep 2018), *evobiR* (Blackmon and Adams 2015), *ggnewscale* (Campitelli 2023), *ggstar* (Xu 2022), *ggtree* (Yu et al. 2017), *ggtreeExtra* (Xu et al. 2021), and *tidytree* (Yu 2022).

Differential expression analyses

First, we trimmed the reads using TrimGalore! (Martin 2011). We then mapped the reads from all three species to both the *Costus allenii* and *Costus lasius* genome (Harenčár et al. 2023). We ran 2pass mapping using STAR (Dobin et al. 2013). In general, we found higher levels of multi-mapped *C. allenii* reads when mapped to *C. allenii*. To investigate this we used mmquant (with merge, -m, set to 10) to count the reads, a tool which can account for multi-mapping reads (Zytnicki 2017). The output from mmquant was read into R for further analysis. To calculate percentage of gene expression due to TPS expression, we normalized reads by library size using TMM (trimmed mean of M values) (Robinson and Oshlack 2010). Due to the consistency in TPS expression as calculated by mapping to both genomes, we decided to use the *C. lasius* genome assembly to identify genes differentially expressed between the three species to remove the complication of multi-mapping. To do this, we first counted reads using featureCounts (Liao et al. 2014) and again read the output into R. We filtered genes to include those with at least one count per million (CPM) in at least one library and normalized reads by library size using TMM (trimmed mean of M values). Differential gene expression was evaluated using the non-parametric "noiseqbio" function from the *NOISeq* package (Tarazona et al. 2011, 2015). We

used the default cut-off of q=0.95 in the “degenes” function which is equivalent to an adjusted p-value of 0.05.

GO Analysis

We functionally annotated *C. alennii* genes using Blast2GO (Götz et al. 2008; Burge et al. 2012). We imported these functional annotations into R to carry out enrichment analysis with the package *topGO* (Alexa and Rahnenfuhrer 2022). We considered the background to be all genes included in that specific differential expression analysis after filtering.

Plotting and data manipulation

Additional packages used for general plotting: *cowplot* (Wilke 2020), *ggplot2* (Wickham 2009), and *scico* (Pedersen and Crameri 2023). Packages used for data transformation and manipulation include: *data.table* (Barrett et al. 2024), *dplyr* (Wickham et al. 2021), *knitr* (Xie 2023), *magrittr* (Bache and Wickham 2023), *readr* (Wickham et al. 2023), *reshape2* (Wickham 2007), and *tibble* (Müller and Wickham 2022). Analyses were carried out in R version 4.3.2 (R Core Team 2023).

Data availability

Data and R scripts used for analysis are available from Open Science Framework: https://osf.io/2ap4k/?view_only=6b8f5ff7d2ae499b871e1b1fb6fe6219. The raw sequence reads and final genome assembly for *Costus allenii* are deposited in the SRA (BioProject PRJNA1091680) https://dataview.ncbi.nlm.nih.gov/object/PRJNA1091680?reviewer=p4jo1ioj3vqualqk1kvege7aql. The raw RNAseq data is deposited in the SRA (BioProject PRJNA1010186) https://dataview.ncbi.nlm.nih.gov/object/PRJNA1010186?reviewer=fil9vpq6up5g3pj7e654rtokje

Acknowledgements

This work was funded by the National Science Foundation Dimensions of Biodiversity

grant awarded to K.M.K. (DEB-1737889) and S.R.R. (DEB-1737771), and the Jean H. Langenheim Chair in Plant Ecology and Evolution held by K.M.K. We thank the Organization for Tropical Studies, Estación Biológica Monteverde, and Estación Tropical La Gamba for facilitating field work in Costa Rica and the Smithsonian Tropical Research Institute for facilitating field work in Panama. All field research was conducted with appropriate research permits in Costa Rica (M-P-SINAC-PNI-ACAT-026-2018, ACC-PI-0272018, MPC-SINAC-PNI-ACLAP-020-2018, INV-ACOSA-076-18, MPC-SINAC-PNI-ACTo-020-18, CONAGEBIO R-056-2019OT and R-058-2019-OT) and Panama (SE/AP-13-19, 09-99) We thank Julia Harenčár for assistance with lab work, Pedro Juarez and Cecilia Girvin for assistance with field sampling, and Alyssa Kaatmann, Aubrie Tait, Kate Uckele, Melina Sauerman Carrizosa, and Selena Vengco for assistance with greenhouse sampling. We thank J. Velzy and S. Childress for greenhouse care of the UCSC living *Costus* collection. We thank the Ramírez lab for helpful discussions.

Author Contributions

Conceptualization, K.D., K.M.K., and S.R.R.; Investigation, K.D. and K.M.K.; Formal Analysis, K.D.; Visualization, K.D.; Resources, K.M.K. and S.R.R.; Writing – Original Draft, K.D.; Writing – Review and Editing, K.D., K.M.K., and S.R.R.; Supervision, K.M.K. and S.R.R.; Project Administration, K.M.K. and S.R.R.; Funding Acquisition, K.M.K. and S.R.R.

Declaration of Interests

The authors declare no competing interests.

References

Abrahamczyk, S., and S. S. Renner. 2015. The temporal build-up of hummingbird/plant mutualisms in North America and temperate South America. BMC Evol. Biol. 15:104.

Adams, D. C. 2014. A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. Syst. Biol. 63:685–697.

Adams, D. C., and M. L. Collyer. 2018. Phylogenetic ANOVA: Group-clade aggregation, biological challenges, and a refined permutation procedure. Evolution 72:1204–1215.

Adams, D. C., and E. Otárola-Castillo. 2013. geomorph: an r package for the collection and analysis of geometric morphometric shape data. Methods Ecol. Evol. 4:393–399.

Albert, V. A., S. E. Williams, and M. W. Chase. 1992. Carnivorous Plants: Phylogeny and Structural Evolution. Science 257:1491–1495. American Association for the Advancement of Science.

Alexa, A., and J. Rahnenfuhrer. 2022. topGO: Enrichment Analysis for Gene Ontology.

Alonge, M., L. Lebeigle, M. Kirsche, K. Jenike, S. Ou, S. Aganezov, X. Wang, Z. B. Lippman, M. C. Schatz, and S. Soyk. 2022. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. Genome Biol. 23:258.

Amrad, A., M. Moser, T. Mandel, M. de Vries, R. C. Schuurink, L. Freitas, and C. Kuhlemeier. 2016. Gain and Loss of Floral Scent Production through Changes in Structural Genes during Pollinator-Mediated Speciation. Curr. Biol. 26:3303–3312.

Arendsee, Z. 2017. rhmmer: utilities parsing “HMMER” results.

Bache, S. M., and H. Wickham. 2023. magrittr: A Forward-Pipe Operator for R.

Barreto, E., M. M. A. Boehm, E. Ogutcen, S. Abrahamczyk, M. Kessler, J. Bascompte, A. S. Dellinger, C. Bello, D. M. Dehling, F. Duchenne, M. Kaehler, L. P. Lagomarsino, L. G.

Lohmann, M. A. Maglianesi, H. Morlon, N. Muchhala, J. F. Ornelas, M. Perret, N. R. Salinas, S. D. Smith, J. C. Vamosi, I. G. Varassin, and C. H. Graham. 2024. Macroevolution of the plant–hummingbird pollination system. Biol. Rev. 000.

Barrett, T., M. Dowle, A. Srinivasan, J. Gorecki, M. Chirico, and T. Hocking. 2024. data.table: Extension of `data.frame`.

Bergamo, P. J., A. R. Rech, V. L. G. Brito, and M. Sazima. 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.

Blackmon, H., and R. A. Adams. 2015. EvobiR: Tools for comparative analyses and teaching evolutionary biology. , doi: 10.5281/zenodo.30938.

Burge, S., E. Kelly, D. Lonsdale, P. Mutowo-Muellenet, C. McAnulla, A. Mitchell, A. Sangrador-Vegas, S.-Y. Yong, N. Mulder, and S. Hunter. 2012. Manual GO annotation of predictive protein signatures: the InterPro approach to GO curation. Database J. Biol. Databases Curation 2012:bar068.

Byers, K. J. R. P., H. D. Bradshaw, and J. A. Riffell. 2014. Three floral volatiles contribute to differential pollinator attraction in monkeyflowers (*Mimulus*). J. Exp. Biol. 217:614–623.

Cabanettes, F., and C. Klopp. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ 6:e4958. PeerJ Inc.

Campitelli, E. 2023. ggnewscale: Multiple Fill and Colour Scales in "ggplot2."

Capella-Gutierrez, S., J. M. Silla-Martinez, and T. Gabaldon. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.

Castellanos, M. C., P. Wilson, and J. D. Thomson. 2004. ‘Anti-bee’ and ‘pro-bird’ changes during the evolution of hummingbird pollination in *Penstemon* flowers. J. Evol. Biol. 17:876–885.

Chen, G. F., and D. W. Schemske. 2015. Ecological differentiation and local adaptation in two sister species of Neotropical *Costus* (Costaceae). Ecology 96:440–449.

Chen, L., A. L. DeVries, and C.-H. C. Cheng. 1997. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. Proc. Natl. Acad. Sci. 94:3817–3822. Proceedings of the National Academy of Sciences.

Cheng, H., G. T. Concepcion, X. Feng, H. Zhang, and H. Li. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat. Methods 18:170–175. Nature Publishing Group.

Collyer, M. L., and D. C. Adams. 2018. RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. Methods Ecol. Evol. 9:1772–1779.

Dellinger, A. S., O. Paun, J. Baar, E. M. Temsch, D. Fernández-Fernández, and J. Schönenberger. 2022. Population structure in Neotropical plants: Integrating pollination biology, topography and climatic niches. Mol. Ecol. 31:2264–2280.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21.

Driver, R. J. 2022. Evolution of olfactory receptors in birds. East Carolina University.

Driver, R. J., and C. N. Balakrishnan. 2021. Highly Contiguous Genomes Improve the Understanding of Avian Olfactory Receptor Repertoires. Integr. Comp. Biol. 61:1281–1290.

Eddy, S. R. 2011. Accelerated Profile HMM Searches. PLOS Comput. Biol. 7:e1002195.

Farré-Armengol, G., I. Filella, J. Llusià, and J. Peñuelas. 2017. β-Ocimene, a Key Floral and Foliar Volatile Involved in Multiple Interactions between Plants and Other Organisms. Molecules 22:1148.

Fernandes, N. de S., F. A. N. Silva, F. A. S. de Aragão, G. J. Zocolo, and B. M. Freitas. 2019. Volatile Organic Compounds Role in Selective Pollinator Visits to Commercial Melon Types. J. Agric. Sci. 11:p93.

Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152.

Gamba, D., and N. Muchhala. 2023. Pollinator type strongly impacts gene flow within and among plant populations for six Neotropical species. Ecology 104:e3845.

Gershenzon, J., and N. Dudareva. 2007. The function of terpene natural products in the natural world. Nat. Chem. Biol. 3:408–414.

Goldsmith, K. M., and T. H. Goldsmith. 1982. Sense of Smell in the Black-Chinned Hummingbird. The Condor 84:237.

Götz, S., J. M. García-Gómez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talón, J. Dopazo, and A. Conesa. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36:3420–3435.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644–652. Nature Publishing Group.

Graham, A. M., J. M. Jamison, M. Bustos, C. Cournoyer, A. Michaels, J. S. Presnell, R. Richter, D. E. Crocker, A. Fustukjian, M. E. Hunter, L. D. Rea, J. Marsillach, C. E. Furlong, W. K. Meyer, and N. L. Clark. 2023. Reduction of Paraoxonase Expression Followed by Inactivation across Independent Semiaquatic Mammals Suggests Stepwise Path to Pseudogenization. Mol. Biol. Evol. 40:msad104.

Grant, K. A., and V. Grant. 1968. Hummingbirds and their flowers. Columbia University Press, New York.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman, and A. Regev. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8:1494–1512. Nature Publishing Group.

Harenčár, J., O. M. Vargas, M. Escalona, D. W. Schemske, and K. M. Kay. 2023. Genome assemblies and comparison of two neotropical spiral gingers: *Costus pulverulentus* and *C. lasius*. J. Hered. esad018.

Hervé, M. 2021. RVAideMemoire: Testing and Plotting Procedures for Biostatistics. R package version 0.9-81.

Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol. Biol. Evol. 35:518–522. Oxford Academic.

Janzen, D. H. 1981. Bee Arrival at Two Costa Rican Female Catasetum Orchid Inflorescences, and a Hypothesis on Euglossine Population Structure. Oikos 36:177–183.

Janzen, D. H. 1971. Euglossine Bees as Long-Distance Pollinators of Tropical Plants. Science 171:203–205.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14:587–589.

Katoh, K., and D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30:772–780. Oxford Academic.

Kay, K. M., and D. L. Grossenbacher. 2022. Evolutionary convergence on hummingbird pollination in Neotropical *Costus* provides insight into the causes of pollinator shifts. New Phytol. 236:1572–1583.

Kay, K. M., and D. W. Schemske. 2003. Pollinator Assemblages and Visitation Rates for 11 Species of Neotropical *Costus* (Costaceae). Biotropica 35:198–207.

Knudsen, J. T., L. Tollsten, I. Groth, G. Bergstöm, and R. A. Raguso. 2004. Trends in floral scent chemistry in pollination syndromes: floral scent composition in hummingbird-pollinated taxa. Bot. J. Linn. Soc. 146:191–199.

Krauss, S. L., R. D. Phillips, J. D. Karron, S. D. Johnson, D. G. Roberts, and S. D. Hopper. 2017. Novel Consequences of Bird Pollination for Plant Mating. Trends Plant Sci. 22:395–410.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359. Nature Publishing Group.

Lanier, E. R., T. B. Andersen, and B. Hamberger. 2023. Plant terpene specialized metabolism: complex networks or simple linear pathways? Plant J. 114:1178–1201.

Larue, A.-A. C., R. A. Raguso, and R. R. Junker. 2016. Experimental manipulation of floral scent bouquets restructures flower–visitor interactions in the field. J. Anim. Ecol. 85:396–408.

Li, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100.

Liao, Y., G. K. Smyth, and W. Shi. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinforma. Oxf. Engl. 30:923–930.

Liu, J. W., P. Milet-Pinheiro, G. Gerlach, M. Ayasse, C. E. P. Nunes, I. Alves-Dos-Santos, and S. R. Ramírez. 2024. Macroevolution of floral scent chemistry across radiations of male euglossine bee-pollinated plants. Evolution 78:98–110.

Losos, J. B. 2011. CONVERGENCE, ADAPTATION, AND CONSTRAINT. Evolution 65:1827–1840.

Maas, P. J. M. 1972. Costoideae (Zingiberaceae). Flora Neotropica 8:1–139. New York Botanical Garden Press.

Maas, P. J. M. 1977. Renealmia (Zingiberaceae - Zingiberoideae) Costoideae (Additions) (Zingiberaceae). Flora Neotropica 18:1–218. New York Botanical Garden Press.

Manni, M., M. R. Berkeley, M. Seppey, F. A. Simão, and E. M. Zdobnov. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol. Biol. Evol. 38:4647–4654.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10–12.

Moran, D., R. Softley, and E. J. Warrant. 2015. The energetic cost of vision and the evolution of eyeless Mexican cavefish. Sci. Adv. 1:e1500363. American Association for the Advancement of Science.

Müller, K., and H. Wickham. 2022. tibble: Simple Data Frames.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol. Biol. Evol. 32:268–274.

Oksanen, J., F. Guillaume Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. Minchin, R. O'Hara, G. Simpson, P. Solymos, H. Stevens, E. Szoecs, and H. Wagner. 2020. vegan: Community Ecology Package. R package version 2.5-7.

Opedal, Ø. H., M. Falahati-Anbaran, E. Albertsen, W. S. Armbruster, R. Pérez-Barrales, H. K. Stenøien, and C. Pélabon. 2017. Euglossine bees mediate only limited long-distance gene flow in a tropical vine. New Phytol. 213:1898–1908.

Orr, H. A. 2005. The Probability of Parallel Evolution. Evolution 59:216–220.

Paradis, E., and K. Schliep. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, doi: 10.1093/bioinformatics/bty633.

Pedersen, T. L., and F. Crameri. 2023. scico: Colour Palettes Based on the Scientific Colour-Maps.

Peng, F., K. J. R. P. Byers, and H. D. Bradshaw Jr. 2017. Less is more: Independent loss-of-function OCIMENE SYNTHASE alleles parallel pollination syndrome diversification in monkeyflowers (*Mimulus*). Am. J. Bot. 104:1055–1059.

Petrén, H., T. G. Köllner, and R. R. Junker. 2023. Quantifying chemodiversity considering biochemical and structural properties of compounds with the R package chemodiv. New Phytol. 237:2478–2492.

Pichersky, E., and R. A. Raguso. 2018. Why do plants produce so many terpenoid compounds? New Phytol. 220:692–702.

Powers, J., K. Eisen, D. Campbell, and R. Raguso. 2023. bouquet: Filter and Augment Biological GC-MS Data.

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raguso, R. A. 2016. Plant Evolution: Repeated Loss of Floral Scent — A Path of Least Resistance? Curr. Biol. 26:R1282–R1285.

Ramos, S. E., and F. P. Schiestl. 2019. Rapid plant evolution driven by the interaction of pollination and herbivory. Science 364:193–196. American Association for the Advancement of Science.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

Robinson, M. D., and A. Oshlack. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 11:R25.

Rosas-Guerrero, V., R. Aguilar, S. Martén-Rodríguez, L. Ashworth, M. Lopezaraiza-Mikel, J. M. Bastida, and M. Quesada. 2014. A quantitative review of pollination syndromes: do floral traits predict effective pollinators? Ecol. Lett. 17:388–400.

Sackton, T. B., and N. Clark. 2019. Convergent evolution in the genomics era: new insights and directions. Philos. Trans. R. Soc. B Biol. Sci. 374:20190102. Royal Society.

Sas, C., F. Müller, C. Kappel, T. V. Kent, S. I. Wright, M. Hilker, and M. Lenhard. 2016. Repeated Inactivation of the First Committed Enzyme Underlies the Loss of Benzaldehyde Emission after the Selfing Transition in *Capsella*. Curr. Biol. 26:3313–3319.

Schemske, D. W., and H. D. Bradshaw. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). Proc. Natl. Acad. Sci. 96:11910–11915. Proceedings of the National Academy of Sciences.

Shumate, A., and S. L. Salzberg. 2021. Liftoff: accurate mapping of gene annotations. Bioinformatics 37:1639–1643.

Sifuentes-Romero, I., A. M. Aviles, J. L. Carter, A. Chan-Pong, A. Clarke, P. Crotty, D. Engstrom, P. Meka, A. Perez, R. Perez, C. Phelan, T. Sharrard, M. I. Smirnova, A. J. Wade, and J. E. Kowalko. 2023. Trait Loss in Evolution: What Cavefish Have Taught Us about Mechanisms Underlying Eye Regression. Integr. Comp. Biol. 63:393–406.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212.

Slater, G. S. C., and E. Birney. 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31.

Smid, M., R. R. J. Coebergh van den Braak, H. J. G. van de Werken, J. van Riet, A. van Galen, V. de Weerd, M. van der Vlugt-Daane, S. I. Bril, Z. S. Lalmahomed, W. P. Kloosterman, S. M. Wilting, J. A. Foekens, J. N. M. IJzermans, P. P. L. O. Coene, J. W. T. Dekker, D. D. E. Zimmerman, G. W. M. Tetteroo, W. J. Vles, W. W. Vrijland, R. Torenbeek, M. Kliffen, J. H. Carel Meijer, A. A. vd Wurff, J. W. M. Martens, A. M. Sieuwerts, and on behalf of the MATCH study group. 2018. Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. BMC Bioinformatics 19:236.

Tarazona, S., P. Furió-Tarí, D. Turrà, A. D. Pietro, M. J. Nueda, A. Ferrer, and A. Conesa. 2015. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res. 43:e140.

Tarazona, S., F. García-Alcalde, J. Dopazo, A. Ferrer, and A. Conesa. 2011. Differential expression in RNA-seq: a matter of depth. Genome Res. 21:2213–2223.

Tholl, D. 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr. Opin. Plant Biol. 9:297–304.

Thomson, J. D., and P. Wilson. 2008. Explaining Evolutionary Shifts between Bee and Hummingbird Pollination: Convergence, Divergence, and Directionality. Int. J. Plant Sci. 169:23–38. The University of Chicago Press.

Tsuchimatsu, T., and S. Fujii. 2022. The selfing syndrome and beyond: diverse evolutionary consequences of mating system transitions in plants. Philos. Trans. R. Soc. B Biol. Sci. 377:20200510.

Valderrama, E., J. B. Landis, D. Skinner, P. J. M. Maas, H. Maas-van de Kramer, T. André, N. Grunder, C. Sass, M. Pinilla-Vargas, C. J. Guan, H. R. Phillips, A. M. R. de Almeida, and

C. D. Specht. 2022. The genetic mechanisms underlying the convergent evolution of pollination syndromes in the Neotropical radiation of *Costus* L. Front. Plant Sci. 13:874322.

Vargas, O. M., B. Goldston, D. L. Grossenbacher, and K. M. Kay. 2020. Patterns of speciation are similar across mountainous and lowland regions for a Neotropical plant radiation (Costaceae: *Costus*). Evolution 74:2644–2661.

Wessinger, C. A. 2024. How the switch to hummingbird pollination has greatly contributed to our understanding of evolutionary processes. New Phytol. 241:59–64.

Wickham, H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H. 2007. Reshaping Data with the reshape Package. J. Stat. Softw. 21:1–20.

Wickham, H., R. François, L. Henry, and K. Müller. 2021. dplyr: A Grammar of Data Manipulation. R package version 1.0.7.

Wickham, H., J. Hester, and J. Bryan. 2023. readr: Read Rectangular Text Data.

Wikelski, M., J. Moxley, A. Eaton-Mordas, M. M. López-Uribe, R. Holland, D. Moskowitz, D. W. Roubik, and R. Kays. 2010. Large-Range Movements of Neotropical Orchid Bees Observed via Radio Telemetry. PLOS ONE 5:e10738.

Wilke, C. O. 2020. cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2.”

Workman, R., R. Fedak, D. Kilburn, S. Hao, K. Liu, and W. Timp. 2019. High Molecular Weight DNA Extraction from Recalcitrant Plant Species for Third Generation Sequencing.

Wozniak, N., K. Sartori, C. Kappel, L. Zhao, A. Erban, I. Fehrle, F. Jantzen, M. Orsucci, S. Rosa, M. Lenhard, J. Kopka, and A. Sicard. 2022. Convergence and molecular evolution of floral fragrance after independent transitions to self–fertilization. bioRxiv.

Xie, Y. 2023. knitr: A General-Purpose Package for Dynamic Report Generation in R.

Xu, S. 2022. ggstar: Multiple Geometric Shape Point Layer for “ggplot2.”

Xu, S., Z. Dai, P. Guo, X. Fu, S. Liu, L. Zhou, W. Tang, T. Feng, M. Chen, L. Zhan, T. Wu, E. Hu, Y. Jiang, X. Bo, and G. Yu. 2021. ggtreeExtra: Compact Visualization of Richly Annotated Phylogenetic Data. Mol. Biol. Evol. 38:4039–4042.

Yost, J. M., and K. M. Kay. 2009. The evolution of postpollination reproductive isolation in *Costus*. Sex. Plant Reprod. 22:247–255.

Yu, G. 2022. Data Integration, Manipulation and Visualization of Phylogenetic Trees. Chapman & Hall.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T.-Y. Lam. 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.

Zdobnov, E. M., D. Kuznetsov, F. Tegenfeldt, M. Manni, M. Berkeley, and E. V. Kriventseva. 2021. OrthoDB in 2020: evolutionary and functional annotations of orthologs. Nucleic Acids Res. 49:D389–D393.

Zu, P., F. P. Schiestl, D. Gervasi, X. Li, D. Runcie, and F. Guillaume. 2020. Floral signals evolve in a predictable way under artificial and pollinator selection in *Brassica rapa*. BMC Evol. Biol. 20:127.

Zytnicki, M. 2017. mmquant: how to count multi-mapping reads? BMC Bioinformatics 18:411.