

SCIENTIFIC REPORTS



OPEN

Multiple Geographical Origins of Environmental Sex Determination enhanced the diversification of Darwin's Favourite Orchids

Oscar Alejandro Pérez-Escobar¹, Guillaume Chomicki², Fabien L. Condamine³, Jurriaan M. de Vos^{4,7}, Aline C. Martins⁵, Eric C. Smidt⁵, Bente Klitgård¹, Günter Gerlach⁶ & Jochen Heinrichs⁸

Environmental sex determination (ESD) — a change in sexual function during an individual life span driven by environmental cues — is an exceedingly rare sexual system among angiosperms. Because ESD can directly affect reproduction success, it could influence diversification rate as compared with lineages that have alternative reproductive systems. Here we test this hypothesis using a solid phylogenetic framework of Neotropical Catsetinae, the angiosperm lineage richest in taxa with ESD. We assess whether gains of ESD are associated with higher diversification rates compared to lineages with alternative systems while considering additional traits known to positively affect diversification rates in orchids. We found that ESD has evolved asynchronously three times during the last ~5 Myr. Lineages with ESD have consistently higher diversification rates than related lineages with other sexual systems. Habitat fragmentation due to mega-wetlands extinction, and climate instability are suggested as the driving forces for ESD evolution.

Angiosperms have evolved a rich array of sexual systems that are unevenly distributed across lineages^{1,2}. About 90% of angiosperms produce male and female reproductive organs in the same flower^{2–4}. A small fraction of genera (i.e. 7%) include dioecious species, which bear sexes on separate individuals⁵. One unusual sexual system is referred to as labile sex expression, and involves changes in sex expression linked to environmental and phenotypic interactions⁶. A regular form of labile sex expression is the recurrent alternation of functionally male and female flower production over reproductive seasons⁵ (e.g. duodichogamy in *Acer*⁶). However, a more extreme form of labile sexual expression is environmental sex determination (ESD), where an individual can change its sex in response to environmental variables such as sunlight and temperature^{3,6,7}. Therefore, plants with ESD can produce male, female or bisexual flowers either on the same individual or in separate individuals depending on specific environmental conditions⁸. ESD is a rare sexual system in angiosperms and has so far been reported for only ca 450 species in three families⁵. Several authors proposed that habitat heterogeneity afforded by fragmentation may provide conditions where ESD would be beneficial^{7–9}. Given the differential demand of resources by male (e.g. pollen production) and female (e.g. seed production) sex functions, ESD would enable the optimization of resource reallocation to either sex function in response to specific site characteristics.

Orchids, with about 25,000 species, are among the most species-rich families of flowering plants on Earth^{10–12}. They are well known for their diversity of reproductive systems¹³, but these remain poorly understood^{14,15}. The vast majority of Orchidaceae produce male and female reproductive organs in the same flower¹⁶. Indeed, orchid

¹Department of Identification and Naming, Royal Botanic Gardens Kew, Richmond, TW9 3AB, UK. ²Department of Plant Sciences, University of Oxford, South Park Road, OX1 3RB, Oxford, United Kingdom. ³CNRS, UMR 5554 Institut de Sciences de l'Evolution (Université de Montpellier), Place Eugène Bataillon, 34095, Montpellier, France. ⁴Department of Comparative Plant and Fungal Biology, Royal Botanic Gardens Kew, Richmond, TW9 3AB, United Kingdom. ⁵Department of Botany, Federal University of Paraná, PB 19031, Curitiba, PR, 81531-980, Brazil. ⁶Botanischer Garten München, Menzinger Straße 67, D-80638, München, Germany. ⁷Present address: Department of Environmental Sciences - Botany, University of Basel, Totengässlein 3, 4051, Basel, Switzerland. ⁸Department für Biologie I, Systematische Botanik und Mykologie, Ludwig-Maximilians-Universität, Menzinger Straße 67, D-80638, München, Germany. Correspondence and requests for materials should be addressed to O.A.P.-E. (email: o.perez-escobar@kew.org) or J.M.d.V. (email: jurriaan.devos@unibas.ch)

flowers exhibit an extreme case of hermaphroditism in the plant kingdom because of the fusion of male and female organs into a gynostemium^{17,18}. An exception to this rule however are the Catasetinae, a Neotropical orchid lineage containing 354 species^{19,20}. Because of their striking sexual systems (i.e. ESD, protandry, adichogamy) and the diversity of pollination syndromes (e.g. male euglossine-bee pollination, oil bee pollination)^{21,22}, Catasetinae have long attracted the attention of botanists and naturalists²³. However, we still know little about the origin and evolution of ESD in this lineage.

Catasetinae includes eight genera, namely *Catasetum*, *Clowesia*, *Cyanaeorchis*, *Cycnoches*, *Dressleria*, *Galeandra*, *Grobya* and *Mormodes*, and most of the diversity is found in Central America and Amazonia regions²⁴. Sex expression in Catasetinae, as in several other plant lineages with ESD, is regulated by ethylene production. The pioneer work of Gregg²⁵ documented the role of sunlight exposure and ethylene synthesis on sex determination in *Cycnoches* and *Catasetum*. She convincingly demonstrated that ethylene concentrations are about two- to 100-fold higher in inflorescence apices that are fully exposed to sunlight compared with those kept in shade. Consequently, inflorescences with higher ethylene concentration usually produce female flowers, whereas those with lower levels produce male flowers²⁵.

To date, about one third of the Catasetinae (i.e. ca 100 species) are known to have ESD. Therefore, the group provides a unique opportunity to study its evolutionary context in a phylogenetic framework. A recent study provided strong support for three independent origins of sexual plasticity in Catasetinae^{14,26}. ESD evolved in *Catasetum*, *Cycnoches* and *Mormodes*, always from a protandrous ancestor. Protandry, in turn, was found to be the ancestral condition in the so-called core Catasetinae (*Catasetum*, *Clowesia*, *Cycnoches*, *Dressleria* and *Mormodes*). ESD could be a switch in the evolution of dioecy (i.e. spatial separation of sexes)²¹, as observed in other plant lineages with sexual plasticity such as *Fuchsia* and *Hebe*^{27–29}.

It is still unclear whether the independent origins of ESD occurred synchronously and under the same environmental conditions in a single biogeographical region or evolved asynchronously in different regions, possibly under different environmental conditions. It also remains elusive whether lineages with ESD are associated with higher diversification rates compared with hermaphroditic lineages. In view of its rarity, we hypothesise that ESD could be a disadvantageous system leading to lower diversification rates. Alternatively, ESD could be evolutionarily unstable and could thus represent a transitional state rather than a stable reproductive system.

The extremely high orchid diversity is thought to be related with the evolution of novel traits and the invasion of novel environments³⁰, possibly confounding analysis of the effects of ESD in orchid diversification. Therefore, we also investigated the impact of epiphytism^{10,31} and male euglossine-bee pollination^{21,30} on the diversification dynamics of Catasetinae. Epiphytism and male euglossine-bee pollination are traits that occur in all ESD species of Catasetinae, but these traits are shared with some Catasetinae species that have alternative reproductive systems²⁴ (see *Methods*). Therefore, it is possible that these traits might have influenced the diversification of the subtribe, yet their specific contribution to diversification rates are still unknown.

The limited knowledge on the evolutionary dynamics of ESD in orchids stems partly from the lack of well-sampled species-level phylogenies⁵. This is particularly true for Catasetinae, in which previous studies did not attain even taxon sampling, thus leaving species-rich clades dramatically underrepresented (e.g. *Catasetum*; see^{32–34}). In the current study, we generate the first densely sampled phylogeny of Catasetinae. We combine data on species distribution and a wide range of comparative phylogenetic methods to reconstruct the evolutionary history of ESD. We then infer the biogeographic history of Catasetinae, test whether their ESD lineages have higher diversification rates than non-ESD lineages, while accounting for the possible contribution of other traits to orchid diversification. We also investigate whether past and modern climate niche preferences of lineages with ESD are divergent from non-ESD lineages.

Results

Phylogeny and molecular dating of Catasetinae. Our DNA matrix included a total of 282 species, of which 132 belonged to Catasetinae (i.e. ~37% of the extant diversity of the subtribe; Table S1). Analyses based on Maximum Likelihood (ML) and Bayesian inference (BI) recovered similar topologies. However, they revealed several conflicting positions of species in the nuclear and plastid trees with bootstrap percentage values (MLBS) ≥ 75 and/or posterior probabilities (PP) ≥ 0.95 (Figs S1 and S2). Figures 1 and S3 show the ML trees inferred from concatenated nuclear and non-conflicting plastid sequences, respectively. Here, after exclusion of conflicting positions from the DNA matrices (see *Methods* and Appendix S1 for a detailed explanation on incongruence handling), almost all backbone nodes of the phylogeny achieved MLBS ≥ 75 and PP > 0.95 support.

Divergence time estimates performed under a Bayesian relaxed clock model yielded a mean Coefficient of Variation (CV) value for branch-specific substitution rates of 0.73, indicating substantial among-branch rate heterogeneity, confirming that the use of a relaxed molecular clock is appropriate. Mean ages and the associated 95% credibility intervals obtained under the relaxed clock model for all nodes are provided in Fig. S4. The dating analyses showed that (i) the Catasetinae and Eulophiinae + Cyrtopodiinae shared a common ancestor in a time interval ranging from the early Miocene to the early Oligocene (25.93 Ma \pm 5 Myrs), and (ii) the diversification of the Catasetinae most recent common ancestor (MRCA) started around 19.5 Ma \pm 4 Myrs during the early Miocene.

Geographical origin of ESD and biogeographical history of Catasetinae. Table S4 provides ML statistics for six biogeographical models as inferred using BIOGEOBEARS. The best fitting model was BayArea, including founder-event speciation (J parameter; $\text{LnL}_{\text{best model}} = -860.3$ vs $\text{LnL}_{\text{second best model}} = -875.5$; $\text{AIC}_{\text{best model}} = 1727$ vs $\text{AIC}_{\text{second best model}} = 1757$; Table S4). This model estimated south-eastern South America as the most likely ancestral area of Catasetinae (Fig. 2; Fig. S5). The same region was supported as the ancestral area of *Cyanaeorchis* and *Grobya*. In contrast, *Galeandra* and *Catasetum* inhabited larger areas between Amazonia + Guiana Shield + south-eastern South America and Amazonia + south-eastern South America,

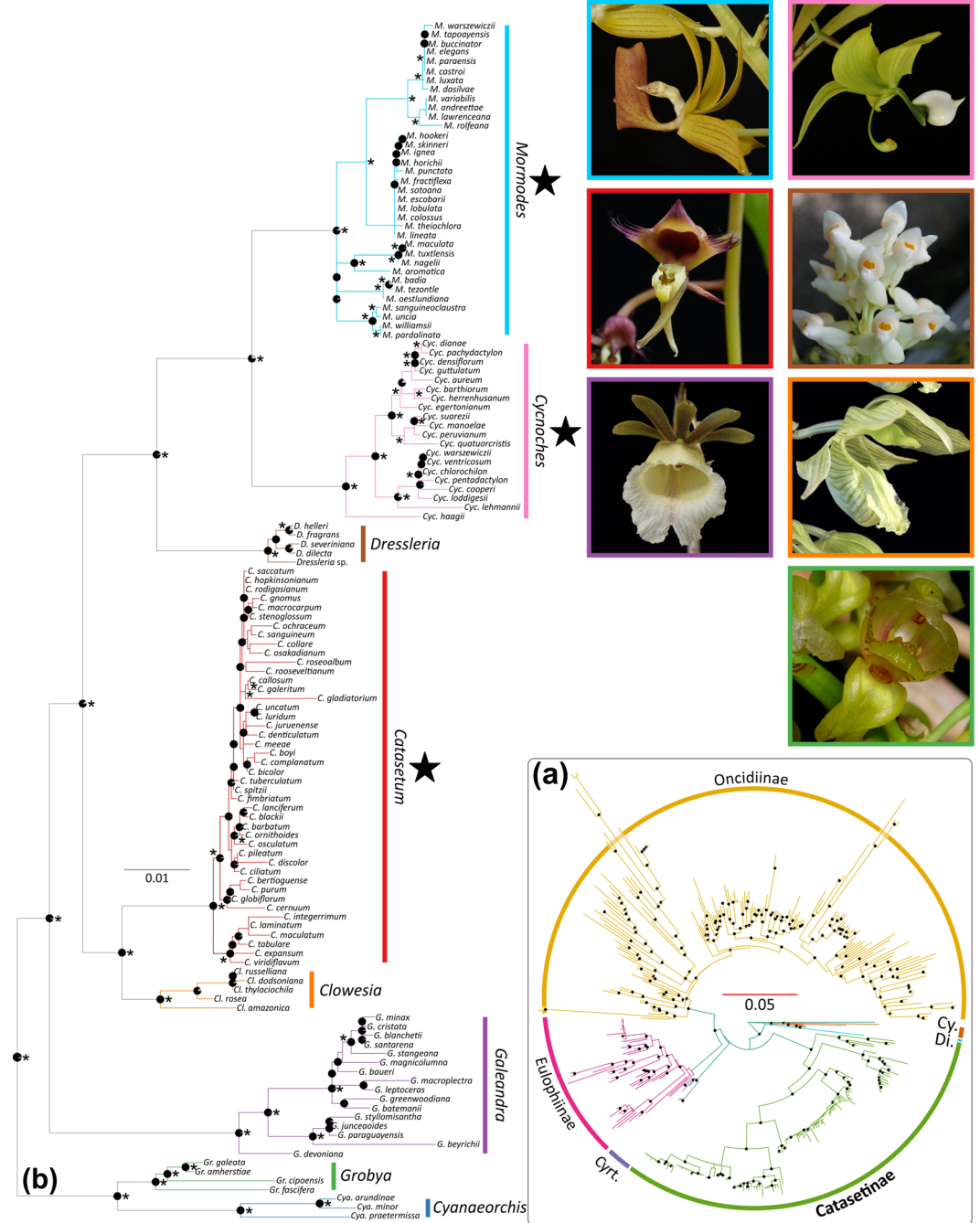


Figure 1. Best scoring ML tree of Catasetinae (b) and sister subtribes (i.e. Eulophiinae, Dipodiinae [Di], Oncidiinae, Cyrtopodiinae [Cyrt.], Cymbidiinae [Cy.]) (a) obtained from non-conflicting concatenated nuclear and plastid loci. Node pie-diagrams indicate Bootstrap percentage values (MLBS > 75), where fully black diagrams indicate MLBS of 100. Asterisks represent BPP values higher than 0.95. Black stars indicate lineages that independently evolved ESD. Representatives of each clade in Catasetinae (except *Cyanaeorchis*) are shown in pictures [inset: *Catasetum*: *Ca. ochraceum*; *Clowesia*: *Cl. russelliana*; *Cycnoches*: *Cyc. lehmannii*; *Dressleria*: *D. eburnea*; *Galeandra*: *Ga. macroplectron*; *Grobya*: *G. galeata*; *Mormodes*: *M. punctata*]. Pictures by: O. Pérez and G. Gerlach.

respectively. The origins of *Clowesia* and *Cycnoches* were estimated to be in Amazonia, whereas the remaining lineages of Catasetinae (i.e. *Dressleria*, *Mormodes*) originated in Central America (Fig. 2; Fig. S5).

Origin, climate niche and evolutionary dynamics of ESD. Ancestral character estimation analysis under the best model (i.e. single transition rate) identified three independent gains of environmental sex determination, without any losses (Fig. 2; Fig. S6). ESD first evolved during the early Pleistocene to late Miocene (5 Ma ± 3 Myrs) at the MRCA of *Cycnoches* in the Amazonia region (Fig. 2; Fig. S5). Subsequently, two gains

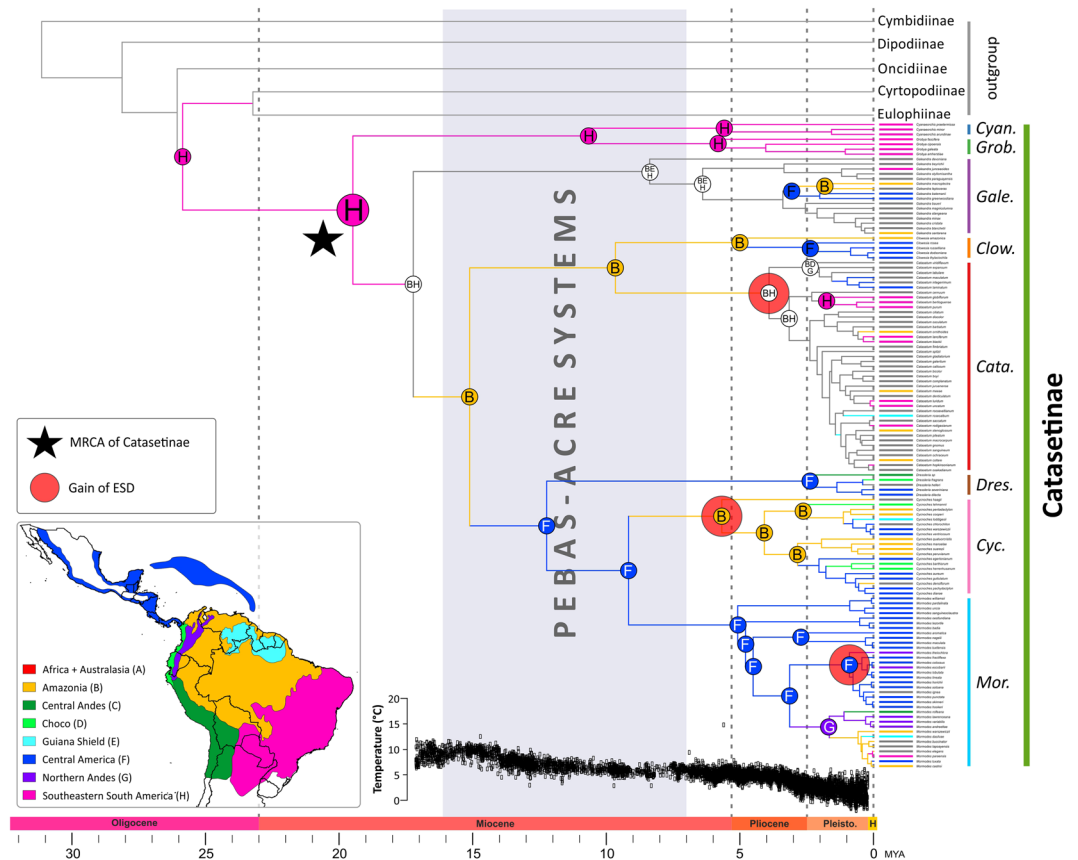


Figure 2. Chronogram of Catasetinae and sister subtribes obtained under a relaxed clock model, applied to a non-conflicting, concatenated nuclear and plastid loci. Minimum and maximum age intervals are provided in Fig. S4. Time scale is provided in million years (Mya). Node charts correspond to ancestral areas estimated under the BAYAREA + J model. The most recent common ancestor (MRCA) of Catasetinae is indicated with a black star, and gains of ESD are indicated with red circles. Coded biogeographical areas are colour-coded following inset map, and are shown in front of taxa names (grey colours indicate taxa distributed in more than one biogeographical region). Catasetinae genera are colour coded as in Fig. 1 (*Mor.*: *Mormodes*; *Cyc.*: *Cycnoches*; *Dres.*: *Dressleria*; *Cata.*: *Catasetum*; *Clow.*: *Clowesia*; *Gale.*: *Galeandra*; *Grob.*: *Grobhya*; *Cyan.*: *Cyanaorchis*). The approximate time span of Pebas and Acre mega-wetlands³⁶ is indicated by a grey rectangle. Inset: Coded areas used for biogeographical analysis. Geopolitical boundaries map generated by ArcMAP (<http://www.sri.com>) using political divisions and elevation data from DIVA-GIS (<http://www.diva-gis.org/data>). Climate curve of the last 17 Myr represented as a function of oxygen-isotope records⁶¹.

of ESD occurred towards the Pleistocene to Pliocene ($3 \text{ Ma} \pm 1 \text{ Myrs}$) and the Pleistocene ($1 \text{ Ma} \pm 1 \text{ Myrs}$) in Amazonia – south-eastern South America and Central America regions, respectively. These gains took place at the MRCAs of *Catasetum* and a small clade of *Mormodes* species restricted to Central America and Northern Andes, respectively (Fig. 2; Figs S5, S6). Lineages with ESD further diversified mostly in lowland areas (i.e. Amazonia, Choco, south-eastern South America, Central America, Guiana Shield). Two episodic colonisations from Central America towards Northern Andes were recovered.

Bioclimatic variables for 1372 georeferenced occurrences of Catasetinae were adequately captured by the first two axes of a principal component analysis (PCA) (47.9% and 23.6%, respectively). The first PCA axis received contributions from all bioclimatic variables except BIO 3 (isothermality), BIO 4 (temperature seasonality), BIO 7 (temperature annual range), BIO 15 (precipitation seasonality), and altitude (Fig. S7). The second PCA axis mostly discriminated dry/hot environments, but also included contributions from BIO 15 (precipitation seasonality; Fig. S8). Therefore, discriminations between particular environments were not detected (Figs S7–S9), as bioclimatic variables representing precipitation and temperature contributed similar amounts of variance to PCA axes 1 and 2 (Figs S7, S8). Overall, the ordination of climatic niche space of lineages based on the bioclimatic variables indicated no evidence for general ecological segregation between lineages with and without ESD (Fig. S9a).

To further assess the apparent lack of climatic niche differentiation between ESD and non-ESD taxa, we excluded all auto-correlated variables (see *Methods*), and performed a non-metric dimensional analysis (NMDS) based on altitude, BIO 15 (precipitation seasonality), BIO 18 (precipitation of warmest quarter) and BIO 19 (precipitation of coldest quarter). Similarly, we found extensive overlaps between climatic niches of ESD and non-ESD lineages (Fig. S9b). However, climatic niche differentiation analysis using present-day bioclimatic variables are not directly informative of the climate niche preferences of ancestral lineages in which ESD arose. Thus,

we performed ancestral character estimation analyses of twenty bioclimatic variables to assess whether ancestors having evolved ESD and those with alternative sexual systems diverged in their climatic niche preferences. We found that ancestral mean values of all bioclimatic variables were consistently similar between ancestors that independently evolved ESD and contemporary lineages with other sexual systems (Fig. 3). More importantly, ancestral mean bioclimatic values of lineages with and without ESD revealed little variation through time (Fig. S10).

Trait-dependent diversification rate analysis supported a model with free speciation rates in lineages with and without ESD ($\text{LnL}_{\text{best model}} = -248.85$; $\text{AIC}_{\text{best model}} = 506.035$ vs. $\text{LnL}_{\text{null model}} = -262.41$; $\text{AIC}_{\text{null model}} = 531.022$, Table 1; Table S5). Under the best model, speciation rates were ~1.6-fold higher in lineages with ESD, while extinction rates were modelled under the same parameter (Fig. S11a; Table 1). We confirmed the robustness of these results through our tailored data simulations (see *Methods*), that generated an expectation about the distribution of ΔAIC values obtained from randomly reshuffled trait data (see *Extended Materials and Methods of Appendix S1*). The distribution of simulated ΔAIC did not overlap empirical AIC values. This suggests that our trait-dependent analysis is not biased by type I error (Fig. S14a).

Evolutionary dynamics of male euglossine-bee pollination and epiphytism. Binary-state dependent diversification analyses of male euglossine-bee pollination and epiphytism supported models with free speciation rates and equal extinction and transition rates (M_2 ; see *Methods*), and free speciation and transition rates but equal extinction rates (M_6), respectively. Both models recovered slightly higher speciation rates for male euglossine-bee pollination (2.13 events/Myr; Fig. S12; Table 1; Table S6), and epiphytism (1.56 events/Myr; Fig. S13; Table 1; Table S7) than for lineages with alternative reproductive systems (1.82 events/Myr; Table S6) and plant habit (1.27 events/Myr; Table S7). Analysis of simulated data confirmed the robustness of real data analysis results for male euglossine-bee pollination. The distribution of ΔAIC values from real data and simulated data did not overlap: $\Delta\text{AIC}_{\text{simulated}}$ was centred towards 30 units while the $\Delta\text{AIC}_{\text{real data}}$ was estimated at 9 (Fig. S14b). Analysis of simulated data for epiphytism did not reach convergence, hence the veracity of binary-state dependent diversification analysis on the real data set could not be evaluated.

Discussion

Forest fragmentation as the context for the multiple geographical origins of ESD. Orchids have a rich variety of reproductive systems³⁵, of which hermaphroditism is the most widespread in the family^{16,35}, reflecting the angiosperm-wide pattern². The Catasetinae, however, have evolved three reproductive systems (i.e. adichogamy, dichogamy, and ESD¹⁴), thus providing a unique opportunity to investigate biotic and abiotic factors shaping the evolution of these sexual systems. By relying on a solid phylogenetic framework, divergence time estimations, and ancestral character estimation, our study demonstrates the independent temporal and geographic origins of ESD in Catasetinae. More importantly, it reveals for the first time an apparent positive correlation between gains of sexual plasticity and higher speciation rates. However, our analyses show that other traits (i.e. male euglossine-bee pollination^{21,30}, and perhaps epiphytism³⁰), may also have positively contributed to the diversification of the Catasetinae.

Ancestral area estimation analyses suggest that sexual plasticity first originated in the Amazonian region during the late Miocene (5 Ma \pm 3 Myrs), a period preceded by the disappearance of large aquatic systems that dominated Amazonia from 17 to 7 Ma³⁶. Such aquatic environments, known as Pebas Lake and Acre Systems, ranged from the drainage of western Amazonia to the Caribbean coast of Venezuela^{37,38} and prompted the fragmentation of pre-existing forests. The disappearance of Pebas and subsequent aquatic systems towards the late Miocene gave rise to flood-plains with a patchy vegetation dominated by palms, ferns and grasses^{38,39}.

The late Miocene vegetation changes in the Amazonian region may also be related to the global cooling subsequent to the mid-Miocene Climatic Optimum (~15 Ma). These climate changes might have promoted forest fragmentation by range expansions and contractions of species due to dramatic temperature changes^{37,40,41}. An example of the effect of drastic temperature changes on vegetation ranges is the peculiar *Campos Rupestres* in Brazil. There, dry periods provided the optimal conditions for vegetation to expand, allowing species to extend their distribution ranges and subsequently occupy larger areas. Succeeding wetter periods would have prompted the contraction of distribution ranges, leading to the formation of isolated vegetation patches^{42,43}. As a result, species like the orchids *Sarcoglottis caudata* and *Liparis beckeri* now occupy very restricted ranges, whereas other taxa have a disjunctive distribution, e.g. *Vellozia virgata*.

Sexual plasticity has been regarded as an adaptation to patchy environments^{7,8,44}, and this might have explanatory power for the gains of sexual plasticity in Catasetinae. The existence of patchy forests towards the end of the late Miocene in South America and vegetation changes in Central America might have facilitated the evolution of ESD by leading individuals to invest different amounts of resources to male and female gamete production. Resource reallocation to either function might thus have taken place as a function of an individual's specific site characteristics^{9,44}.

In a seminal paper, Zimmerman⁴⁵ provided evidence towards the biased sex expression in *Catsetum viridiflavum*. In this species, female plants are more frequently distributed in open canopies whereas male plants often grow in closed canopies. Plants growing in open canopies are longer exposed to sunlight irradiation, thus favouring accumulation of energetic resources (i.e. plant biomass) that are later reallocated to female flower and fruit production⁴⁵. The latter process requires considerable amounts of energy⁴⁶ because of the massive number of seeds produced in a single capsule (see below). Sunlight irradiation appears to be the main ecological variable driving male/female flower production ratio in Catasetinae^{25,45}. However, how other variables tied to resource constrains and patch quality (e.g. substrate type⁴⁵, orchid-mycorrhiza interactions⁴⁷) affect sex ratios in the Catasetinae deserves further research.

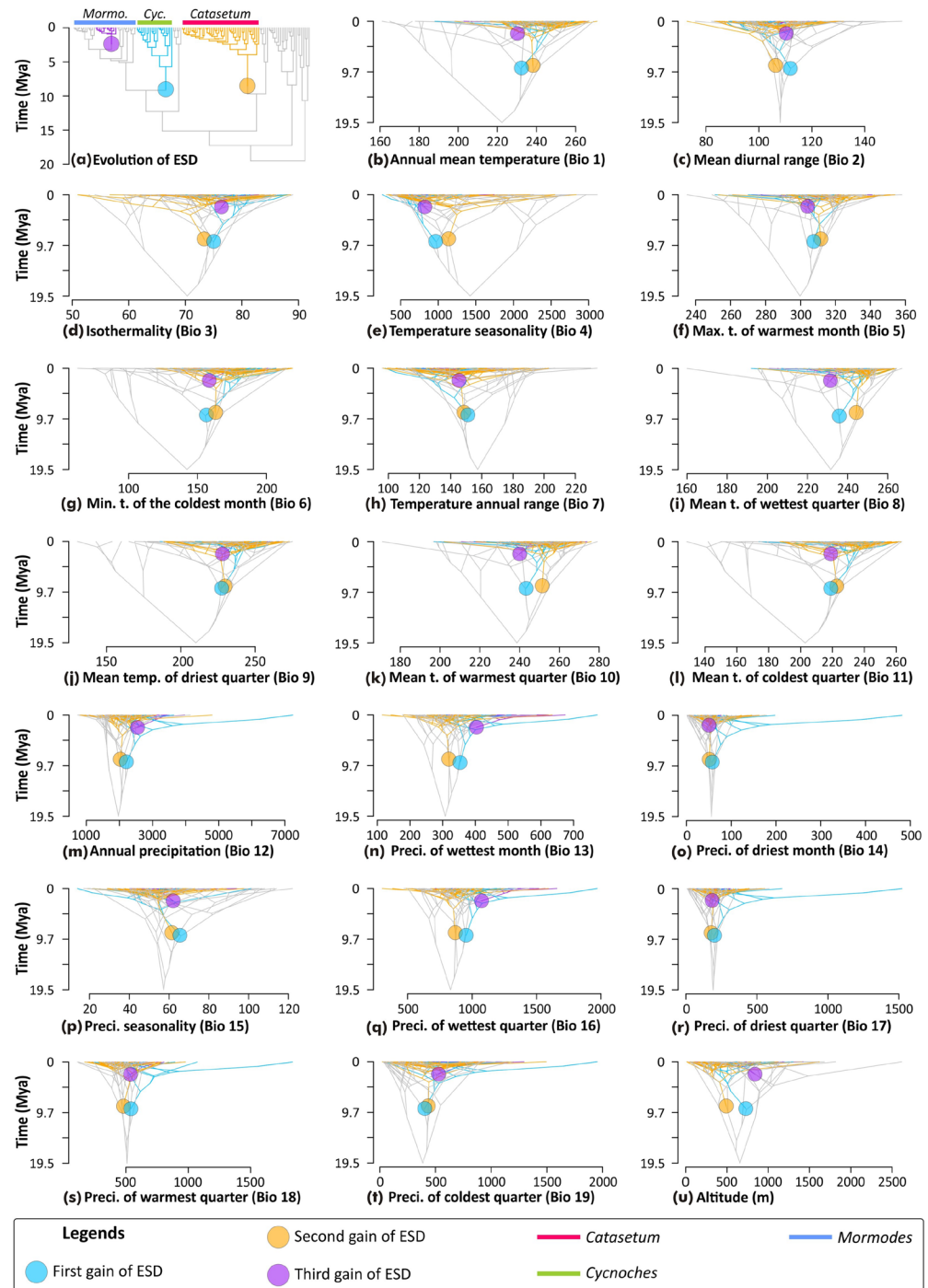


Figure 3. Evolution of ESD and ancestral climatic-niche preferences in Catasetae depicted as traitgrams obtained via ACE analyses of mean values of bioclimatic variables. In all panels, coloured circles denote independent origins of ESD (colour-coded by chronological order). Colour coded nodes and branches denote lineages estimated to have ESD. Grey nodes and branches denote lineages inferred to have evolved alternative sexual systems (i.e. no ESD). Lineages in which ESD has evolved are colour-coded accordingly (*Mormo.*: *Mormodes*; *Cyc.*: *Cycnoches*). (a) Topological position of ESD origins (for posterior probabilities, see Fig. S6) (b) Annual mean temperature (Bio 1); (c) Mean diurnal range (Bio 2); (d) Isothermality (Bio 3); (e) Temperature seasonality (Bio 4); (f) Maximum temperature of warmest month (Bio 5); (g) Maximum temperature of coldest month (Bio 6); (h) Temperature annual range (Bio 7); (i) Mean temperature of wettest quarter (Bio 8); (j) Mean temperature of driest quarter (Bio 9); (k) Mean temperature of warmest quarter (Bio 10); (l) Mean temperature of coldest quarter (Bio 11); (m) Annual precipitation (Bio 12); (n) Precipitation of wettest month (Bio 13); (o) Precipitation of driest month (Bio 14); (p) Precipitation seasonality (Bio 15); (q) Precipitation of wettest quarter (Bio 16); (r) Precipitation of driest quarter (Bio 17); (s) Precipitation of warmest quarter (Bio 18); (t) Precipitation of coldest quarter (Bio 19); (u) Altitude (meters). Note the little divergence through time of ancestral mean bioclimatic variables between lineages having ESD, and those with alternative sexual systems.

Trait	Model	logL ¹	AICc ²	$\lambda^3(0)$	$\lambda(1)$	λ ratio (λ_1/λ_0)	$\mu^4(0)$	$\mu(1)$	q_{0-1} ⁵	q_{1-0} ⁶
ESD	M_2	-248.854	506.035	1.253	2.024	1.614	—	1.181	—	0.0058626
Male euglossine-bee pollination	M_2	-240.045	488.416	1.827	2.138	1.174	—	1.817	—	0.0020853
Epiphytism	M_6	-253.720	517.931	1.276	1.56p	1.2507	—	1.237	0.025	0.0000197

Table 1. Best models and parameter values of BiSSE analysis on lineages with ESD vs. alternative sexual systems, epiphytism vs. other plant habits, and male euglossine-bee pollination vs. other pollination systems. ¹Log Likelihood. ²Akaike Informative Criterion corrected. ³Speciation rates. ⁴Extinction rates. ⁵Transition rates from state 0 to 1. ⁶Transition rates from state 1 to 0.

ESD, epiphytism and male euglossine-bee pollination enhanced diversification in Catasetinae. The overwhelmingly low frequency of lineages with ESD compared to the vast majority of hermaphrodite angiosperms⁵ could suggest that sexual plasticity is a disadvantageous mating system or that it is evolutionary unstable. However, our study shows that Catasetinae clades with ESD have higher speciation rates than those with other sexual systems (i.e. adichogamy, protandry; Fig. S11). Here, the overarching comparison of speciation rates obtained from all analysed traits revealed a greater contribution of ESD evolution to the diversification dynamics of Catasetinae (rates ~1.6-fold higher than in lineages with other sexual systems). In contrast, the evolution of epiphytism (~1.2-fold higher) and male euglossine-bee pollination (~1.17-fold higher) affected diversification in the Catasetinae, but to a lesser extent.

Our ancestral character estimation analysis does not reveal any loss of ESD, suggesting that it is an evolutionary stable system in Catasetinae that has been maintained for up to 5 million years. Nevertheless, diversification in Catasetinae has also been prompted by the gain of traits that were earlier shown to influence diversification rates of Neotropical orchids in general (i.e. male euglossine-bee pollination²¹, epiphytism^{30,31}). Taken together, these findings indicate that rate clade diversification is determined by several, potentially interacting factors³⁰.

Vascular epiphytes are a key component of tropical forests⁴⁸, and orchids account for 68% of the total vascular epiphyte diversity⁴⁹. Epiphytism has evolved at least seven times since the Eocene in Orchidaceae⁵⁰, and might have enhanced diversification by allowing colonization of new, largely unoccupied habitats such as branches and trunks of angiosperms³⁰. In Catasetinae, epiphytism occurs in all members of the three most speciose genera, namely *Catasetum*, *Cycnoches* and *Mormodes*^{24,51} (more than 290 species altogether²⁰). It also occurs in all species of *Grobya*²⁴ (five species), *Clowesia* (seven species) and six of the eleven species of *Galeandra*⁵². Epiphytic species from such clades inhabit a wide array of ecosystems and biogeographical regions^{24,53,54}. Contrastingly, terrestrial species are confined to species-poor clades with restricted distribution such as *Dressleria* (twelve species), *Cyanaeorchis* (three species), and a clade of five species nested in *Galeandra*^{24,52,55}.

Pollination by male euglossine-bee evolved independently three times in Orchidaceae since the early Miocene²¹. It may have accelerated diversification rates by promoting switches in pollinators via small changes in the chemical profile of the floral scent, or alternatively by changes in pollinarium placement sites, hence allowing reproductive isolation³⁰. Such pollination syndromes occur in all species of *Catasetum*, *Clowesia*, *Cycnoches*, *Dressleria* and *Mormodes*, which altogether make up ~85% of total species number in Catasetinae. Alternative pollination syndromes (e.g. oil-bee pollination²², food deception) occur in the remaining, species-poor genera of Catasetinae (i.e. *Cyanaeorchis*, *Galeandra*, *Grobya*). Taken together, our results point to a prominent role of ESD in the evolutionary dynamics of the Catasetinae, evidenced by speciation rates ~1.6-fold higher than in orchids without ESD. Nonetheless, we do not exclude the additional contributions of traits, such as euglossine-bee pollination and epiphytism.

The reliability of our diversification rate results is strongly dependent on the power and hypotheses testing ability of our trait-dependant diversification model (i.e. BiSSE). The BiSSE method is known to be affected by within-clade pseudoreplication⁵⁶ and sample size⁵⁷, issues that are inherent to the number of independent origins of the analysed traits, and the sampling proportion of the clade under study, respectively. Thus, extreme care must be used when interpreting the outcome of trait-dependant diversification analyses⁵⁷. However, by comparing different trait-diversification models with randomly reshuffled coding matrices in our simulations, and analysing other traits known to affect diversification rates in orchids, we have partially accounted for pseudo-replication bias. Indeed, our customized simulations revealed that BiSSE analyses of ESD and male euglossine-bee pollination are not affected by Type I errors. Nonetheless, diversification rate analyses can only account for the effects of the traits here considered, hence we cannot entirely rule out that other variables have additional effects⁵⁶.

ESD, protandry and adichogamy share the same climatic niche in Catasetinae. We have not detected significant differences in the ancestral and modern climate niches between Catasetinae lineages with ESD and other sexual systems (Fig. 3; Figs S9 and S10). Thus, particular temperature or precipitation regimes do not appear to have favoured any of these lineages. Species with ESD that occur often in habitats with longer sun-light exposure compared to species without ESD produce bigger pseudobulbs, female flowers and fruits^{45,58}. These species also produce larger capsules than Catasetinae lineages with adichogamy (e.g. *Cyanaeorchis*, *Dressleria*; Pérez-Escobar, pers. obs.). For instance, the capsules of *Cycnoches chlorochilon* produce on average up to four times more seeds (3,770,000) than fruits of *Cymbidium tracyanum* (850,000), an adichogamous species closely related to Catasetinae⁴⁶. Therefore, under the same climatic niche, Catasetinae lineages with ESD might have produced more offspring than lineages with other sexual systems. However, additional data on fruit size, fruit number, and particularly on seed content for all Catasetinae members are needed to statistically test such assumption.

Occurrence of ESD requires a heterogeneous habitat (e.g. fragmented forest matrix), because under homogeneous ecological conditions (e.g. forest understory) plants with ESD will bear predominantly one sex form, and mating success may thus be dramatically reduced. This highlights the important role of climate instability and its effects on habitat heterogeneity since the Miocene Climate Optimum in the evolution and maintenance of ESD. Nevertheless, it is likely that bioclimatic variables do not represent relevant environmental aspects of the niche in which ESD orchids occurs as compared with non-ESD lineages. Also, the climatic niche analyses pinpoint the explanatory power of variables with maximum variance, hence they will be mostly represented in PCA axes 1 and 2. However, such variables are not necessarily the most relevant and informative for our group of study. Therefore, even though ESD and non-ESD species niches largely overlap, ESD lineages might still be dependent on more specific ecological variables that are not necessarily captured by bioclimatic variables but that may yet be driven by climate (e.g. those operating on forest edge matrices⁵⁹, or affecting vegetation structure).

Conclusions

Environmental Sex Determination (ESD) in Catasetinae has long attracted and even confused⁶⁰ botanists and naturalists²³, and the evolutionary and biogeographic context for the evolution of this remarkable trait remained elusive before our study. Our results show that the evolution of ESD involved multiple gains in three biogeographical regions between late Miocene and late Pleistocene - geological periods that were marked by the extinction of pivotal aquatic environments and mega-wetlands^{36,39} and dramatic fluctuations in temperature⁶¹. Fragmentation of vegetation prompted by the demise of these aquatic systems in South America and by dramatic changes in temperature across the Neotropics might have fostered the evolution of sexual plasticity in Catasetinae. In spite of the extreme rarity of sexual plasticity among the rich array of angiosperm reproductive systems⁵, evolution of ESD in Catasetinae appears to be correlated with higher speciation rates. Other traits such as male euglossine-bee pollination and epiphytism, however, also appear to have positively influenced the spectacular diversity of Catasetinae. This illustrates the difficulty in isolating the effect of one trait on diversification rates from effects of others. The genetic contribution of an individual to future generations is enhanced by ESD because it promotes the development of the sex most suited for specific environmental conditions (i.e. sunlight intensity in Catasetinae^{8,44}). We demonstrate that ESD-bearing Catasetinae inhabit the same climatic niche as their non-ESD relatives. Thus, plants with ESD might produce more offspring over plants having alternative sexual systems because of the increased opportunities for optimizing resource allocation to male and female function at each growing site.

Materials and Methods

Taxon sampling, DNA sequencing and phylogenetic analyses. Species names, geographical origins, voucher specimens, and GenBank accession numbers of sequences included in phylogenetic analyses are provided in Table S1. We built our alignment based on the molecular datasets of Whitten *et al.*³³, Salazar *et al.*⁵¹, Neubig *et al.*⁶², Bone *et al.*⁶³, and Pérez-Escobar *et al.*^{34,64}, and further generated 57 nuclear ribosomal and 25 plastid sequences for poorly sampled lineages such as *Catasetum*, *Dressleria* and *Grobya* (Table S1). Because the precise position of Catasetinae inside the tribe Cymbidiidae was elusive^{33,55}, we also sampled selected representatives of closely related subtribes (i.e. Cyrtopodiinae, Dipodiinae, Eulophiinae, and Oncidiinae). Genomic DNA extraction, amplification, sequencing and alignment procedures are the same as in Irimia *et al.*⁶⁵, Pérez-Escobar *et al.*¹⁴, and Bechteler *et al.*⁶⁶. For the newly sampled taxa, we sequenced nuclear ribosomal external and internal transcribed spacers (ETS and ITS, respectively), a fragment of the *Xdh* gene, and also a ~1500 bp fragment of the plastid gene *ycf1*, and the *trnS-trnG* intergenic spacer. Amplification settings and sequencing primers used for ITS, ETS, *Xdh*, *trnS-trnG*, and *ycf1* are specified in Table S2.

Loci were aligned using MAFFT 7.1⁶⁷ and were further manually checked, yielding a total matrix of 281 taxa and 8104 nucleotides (~24% parsimony informative sites). Congruence between nuclear and plastid datasets was assessed following Pérez-Escobar *et al.*³⁴, and using PACo⁶⁸ (Appendix S1). The procedure is available as a pipeline (<http://www.uv.es/cophylpaco/>) and was also employed to identify operational terminal units (OTUs) from the plastid dataset conflicting with the nuclear dataset (potential outliers detected by PACo are shown in Figs S1 and S2). A detailed explanation on PACo and a rationale on outlier handling is provided in the Extended Materials and Methods section of Appendix S1.

Phylogenetic analyses of separate and concatenated loci were carried out under maximum likelihood (ML) and Bayesian inference (BI), following Pérez-Escobar *et al.*^{14,34}, and Feldberg *et al.*⁶⁹. The best-fitting substitution models for ML and Bayesian analyses were obtained for each data partition using jModelTest v.2.1.6⁷⁰ (Table S3). Phylogenetic inference relied on the ML and BI approaches implemented in RAxML-HPC 8.2.4⁷¹ and MrBayes 3.2.2⁷², respectively, and were carried out on the CIPRES Science Gateway computing facility⁷³. Maximum likelihood bootstrap supports (MLBS) were generated for the ML tree and MLBS $\geq 75\%$ considered as good support, and Bayesian Posterior Probabilities (PP) ≥ 0.95 for the Bayesian majority rule consensus topology were regarded as significant^{74,75}.

Molecular clock dating. Divergence time estimates were conducted using the Bayesian relaxed clock approach of BEAST 2.1.3⁷⁶ with a concatenated nuclear + plastid subset of the data obtained after the PACo analysis. Strict and uncorrelated lognormal molecular clock models, both with pure-birth speciation models as recommended for species level sampling⁷⁷, were compared to explore clock-likeness of the data. For calibrating the relaxed clock model, there are few fossils unambiguously assigned to Orchidaceae⁷⁸, and these are placed to lineages very distantly related to Catasetinae (i.e. *Dendrobium*, *Earina*, both Vandeeae). Recently, a new putative orchid fossil from the Dominican amber (early - middle Miocene, 15–20 Ma) was described and assigned to the tribe Cymbidiidae⁷⁹. However, the author provided no evidence to unambiguously assign such fossil to Cymbidiidae or orchids in general. Secondary calibrations are therefore the only option for dating the tree of Catasetinae. To this end, we relied on the age estimates obtained from a fossil-calibrated global phylogeny of Orchidaceae⁵⁰, which

were set with normally distributed priors that reflected uncertainty in the primary analysis⁷⁶. Settings for absolute age estimation are detailed in the *Extended Materials and Methods* of Appendix S1.

Ancestral areas estimation. Species ranges were coded from the literature^{19,60} and from herbarium specimens housed in the herbaria AMES, COL, F, M, MO, SEL, US, using the R-package SpeciesGeoCoder⁸⁰. Distribution data were also obtained from own field observations. Biogeographical areas were defined considering putative areas of endemism of Catasetinae as well as species distributions observed in other plant lineages such as Rubiaceae³⁷ and Bromeliaceae⁸¹. We divided the geographical range of Catasetinae and outgroup taxa in eight biogeographic regions (see *Extended Materials and Methods* of Appendix S1 for a detailed description of coded areas): (1) Central America; (2) Guiana Shield; (3) Amazonia³⁷; (4) Chocó; (5) Northern Andes; (6) Central Andes; (7) South-eastern South America; and (8) Africa and Australasia. We used the R-package BIOGEOBEARS (Biogeography with Bayesian and Likelihood Evolutionary Analysis in R script⁸²) to estimate ancestral areas on the phylogeny of Catasetinae.

Climatic niche analyses. We mapped georeferenced collection records obtained from floras, GBIF database and herbarium specimens (mean of five specimens per species, maximum number of record per species set to ten), and they represent the known distribution of extant Catasetinae species included in our taxon sampling. To query GBIF database, we relied on the function *occ* of the R-package SPOCC⁸³. We further extracted corresponding values of altitude and 19 climatic variables (30 seconds resolution) reflecting temperature and precipitation regimes from the WorldClim database (available at: <http://www.worldclim.org/current>), using the function *extract* of the R-package RASTER⁸⁴. To characterise the climate-niche occupied by Catasetinae lineages with ESD or other sexual systems, we performed Principal Component Analysis (PCA) on the mean values of the 20 environmental variables of each species, using the default function *princomp* of R (Fig. S7). We further explored the contribution of bioclimatic variables to axes with the most variance by plotting the loadings of every bioclimatic variable onto the axes (Figs S8 and S9).

To avoid spurious results arising from inclusion of correlated variables^{63,85}, we determined the Pearson's correlation coefficients between the variables and altitude and then included only variables with a Pearson's correlation coefficient <0.5, taking a single variable in correlated clusters. This way, we selected altitude, BIO 15 (precipitation seasonality), BIO 18 (precipitation of warmest quarter) and BIO 19 (precipitation of coldest quarter) as a set of maximally uncorrelated variables. We analysed these variables using the R-package VEGAN⁸⁶ to perform non-dimensional metric scaling analyses (NMDS) using the dataset of 1372 georeferenced herbarium specimens.

Ancestral character estimation. We coded for absence (state 0) and presence (state 1) of ESD in 131 species of Catasetinae, plus eight taxa of Cyrtopodiinae (see below). We sampled 68 of the 164 known species with ESD in Catasetinae (i.e. ~41% of the ESD diversity). Information on occurrence of ESD was obtained from the literature^{24,58,87}. We estimated the origin of ESD within Catasetinae using a stochastic character mapping approach implemented on a maximum likelihood framework for ancestral character estimation analysis (ACE) implemented by the function *make.simmap* in the R-package PHYTOOLS⁸⁸. Under this approach, we fitted single (ER), symmetrical (SYM), and asymmetrical character transition rate models (ARD) and performed stochastic mapping on 1000 iterations using the maximum clade credibility tree derived from the BEAST dating analysis (see above). Eight members of the Neotropical genus *Cyrtopodium* were selected as outgroup taxa, and they exhibit hermaphroditic flowers (i.e. no ESD).

To assess whether ancestors with ESD and those with alternative sexual systems differed in their climatic niche preferences back in time, we performed ACE analyses of the same nineteen bioclimatic variables plus altitude, obtained from georeferenced occurrences (see *Climate niche analyses* section of *Methods*). ACE of absolute mean values relied on a Maximum Likelihood framework, and were performed on a phylogeny derived from the character stochastic mapping analysis of ESD (using ER as transition rate model, see *Results*). We then visualized the variation through time of all variables (Fig. 3), using the function *phenogram* of the R-package PHYTOOLS⁸⁸. We also plotted mean ancestral values and their corresponding 95% confidence intervals (CI) estimated for every node of the Catasetinae phylogeny, using the R-package GGLOT2⁸⁹ (Fig. S10).

Trait-dependent diversification analyses. We tested whether lineages with ESD (state 1) had higher speciation rates than hermaphroditic, monoecious clades (state 0). In addition, to tease apart the contribution of biotic variables to Catasetinae diversification that are known to positively influence diversification rates in orchids, we investigated the evolutionary dynamics of male euglossine-bee pollination and epiphytism in the subtribe. The former trait occurs in all species of the genera *Catasetum*, *Clowesia*, *Cynoches*, *Mormodes* and *Dressleria*^{21,24,90}, while the latter occur in all species of *Catasetum*, *Clowesia*, *Cynoches*, *Mormodes*, *Dressleria*²⁴, and part of *Galeandra* species⁵². We relied on the Binary State Speciation and Extinction (BiSSE) model⁹¹ to estimate diversification rates associated with each trait. We used the R-package DIVERSITREE⁹¹ to perform eight diversification models in the maximum likelihood framework on 100 randomly sampled dated trees from the Bayesian divergence times analysis. We tested eight models with different configurations of speciation, extinction, and transition rates between characters. Detailed settings for binary-trait dependent diversification analyses, and simulations to account for type I error biases⁹² are provided in Appendix S1.

References

1. Barrett, S. C. H. The evolution of plant reproductive systems: how often are transitions irreversible? *Proc. R. Soc. B Biol. Sci.* **280**, 1–9 (2013).
2. Barrett, S. C. H. The evolution of plant sexual diversity. *Nat. Rev. Genet.* **3**, 274–284 (2002).
3. Tanurdzic, M. & Banks, J. A. Sex-determining mechanisms in land plants. *Plant Cell* **16**, 61–72 (2004).

4. Bateman, R. M., Hilton, J. & Rudall, P. J. Spatial separation and developmental divergence of male and female reproductive units in gymnosperms, and their relevance to the origin of the angiosperm flower. In *Flowers on the tree of life* (eds Wanntorp, L. & De Craene, L.) 315 (Cambridge University Press, 2011).
5. Renner, S. S. The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* **101**, 1588–1596 (2014).
6. Renner, S. S., Beenken, L., Grimm, G. W., Kocyan, A. & Ricklefs, R. E. The evolution of dioecy, heterodichogamy, and labile sex expression in *Acer*. *Evolution* **61**, 2701–2719 (2007).
7. Charnov, E. L. & Bull, J. When is sex environmentally determined? *Nature* **266**, 828–830 (1977).
8. Korpelainen, H. Labile sex expression in plants. *Biol. Rev. Camb. Philos. Soc.* **73**, 157–180 (1998).
9. Guillon, J., Julliard, R. & Leturque, H. Evolution of habitat-dependent sex allocation in plants: superficially similar to, but intrinsically different from animals. *J. Evol. Biol.* **19**, 500–512 (2006).
10. Dodson, C. H. Why are there so many orchid species? *Lankesteriana* **7**, 99–103 (2003).
11. Fay, M. F. & Chase, M. W. Orchid biology: from Linnaeus via Darwin to the 21st century. *Ann. Bot.* **104**, 359–364 (2009).
12. Bateman, R. M. & Rudall, P. J. The good, the bad and the ugly: using naturally occurring terata to distinguish the possible from the impossible in the orchid floral evolution. *Aliso* **22**, 481–496 (2006).
13. Borba, E. L., Barbosa, A. R., Melo, M. C., Gontijo, S. L. & Oliveira, H. O. Mating systems in the Pleurothallidinae (Orchidaceae): evolutionary and systematic implications. *Lankesteriana* **11**, 207–221 (2011).
14. Pérez-Escobar, O. A., Gottschling, M., Whitten, W. M., Salazar, G. & Gerlach, G. Sex and the Catsetinae (Darwin's favourite orchids). *Mol. Phylogenet. Evol.* **97**, 1–10 (2016).
15. Pérez-Escobar, O. A. *et al.* Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytol.* **215**, 891–905 (2017).
16. Dressler, R. L. *Phylogeny and classification of the orchid family*. **20**, (Dioscorides Press, 1993).
17. Rudall, P. J. & Bateman, R. M. Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. *Biol. Rev.* **77**, 403–441 (2002).
18. Rudall, P. J., Perl, C. D. & Bateman, R. M. Organ homologies in orchid flowers re-interpreted using the Musk Orchid as a model. *PeerJ* **1**, e26 (2013).
19. Romero, G. A. In *Genera Orchidacearum: Vol. 5. Epidendroideae (part two)* (eds Pridgeon, A. M., Cribb, P. J., Chase, M. W. & Rasmussen, F. N.) 11–40 (Oxford University Press, 2009).
20. Chase, M. W. *et al.* An updated classification of Orchidaceae. *Bot. J. Linn. Soc.* **177**, 151–174 (2015).
21. Ramirez, S. R. *et al.* Asynchronous diversification in a specialized plant-pollinator mutualism. *Science* **333**, 1742–1746 (2011).
22. Pansarin, L. M., Castro, M. D. E. M. & Sazima, M. Osmophore and elaiophores of *Grobya amherstiae* (Catsetinae, Orchidaceae) and their relation to pollination. *Bot. J. Linn. Soc.* **159**, 408–415 (2009).
23. Darwin, C. *On the various contrivances by which british and foreign orchids are fertilised by insects*. (Appleton and CO., 1877).
24. Pridgeon, A. M., Cribb, P. J., Chase, M. W. & Rasmussen, F. N. *Genera Orchidacearum: Vol. 5. Epidendroideae (part two)*. (Oxford University Press, 2009).
25. Gregg, K. B. Sunlight-enhanced ethylene evolution by developing inflorescences of *Catsetum* and *Cycnoches* and its relation to female flower production. *Bot. Gaz.* **143**, 466–475 (1982).
26. Pérez-Escobar, O. A. Molecular phylogenetics, evolution of sexual systems and historical biogeography of Darwin's favorite orchids (Catsetinae) and Swan orchids (*Cycnoches* Lindl.). (Ludwig-Maximilians Universität, 2016).
27. Atsatt, P. R. & Rundel, P. W. Pollinator maintenance vs. fruit production: partitioned reproductive effort in subdioecious *Fuchsia lycioides*. *Ann. Missouri Bot. Gard.* **69**, 199–208 (1982).
28. Delph, L. F. Sex-differential resource allocation patterns in the subdioecious shrub *Hebe subalpina*. *Ecology* **71**, 1342–1351 (1990).
29. Delph, L. F. & Wolf, D. E. Evolutionary consequences of gender plasticity in genetically dimorphic breeding systems. *New Phytol.* **166**, 119–128 (2004).
30. Givnish, T. J. *et al.* Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. R. Soc. B Biol. Sci.* **282**, 20151553 (2015).
31. Gravendeel, B., Smithson, A., Slik, F. J. W. & Schuiteman, A. Epiphytism and pollinator specialization: drivers for orchid diversity? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**, 1523–1535 (2004).
32. Romero-González, G. A., Batista, J. & Bem, L. A. Synopsis of the genus *Cyrtopodium* (Catsetinae: Orchidaceae). *Harvard Pap. Bot.* **13**, 189–206 (2008).
33. Whitten, W. M., Neubig, K. M. & Williams, N. H. Generic and subtribal relationships in Neotropical Cymbidieae (Orchidaceae) based on *matK/ycf1* plastid data. *Lankesteriana* **13**, 375–392 (2014).
34. Pérez-Escobar, O. A., Balbuena, J. A. & Gottschling, M. Rumbling orchids: how to assess divergent evolution between chloroplast endosymbionts and the nuclear host. *Syst. Biol.* **65**, 51–65 (2016).
35. Tremblay, R. L., Ackerman, J. D., Zimmerman, J. K. & Calvo, R. N. Variation in sexual reproduction in orchids and its evolutionary consequences: a spasmodic journey to diversification. *Biol. J. Linn. Soc.* **84**, 1–54 (2005).
36. Hoorn, C. *et al.* Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330**, 927–931 (2010).
37. Antonelli, A., Nylander, J. A. A., Persson, C. & Sanmartín, I. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proc. Natl. Acad. Sci. USA* **106**, 9749–9754 (2009).
38. Antonelli, A. & Sanmartín, I. Why are there so many plant species in the Neotropics? *Taxon* **60**, 403–414 (2011).
39. Hoorn, C. An environmental reconstruction of the palaeo-Amazon River system (Middle–Late Miocene, NW Amazonia). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **112**, 187–238 (1994).
40. Gentry, A. H. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Ann. Missouri Bot. Gard.* **69**, 557–593 (1982).
41. Pokorný, L. *et al.* Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Front. Genet.* **6**, 1–15 (2015).
42. Alves, R. V. & Kolbek, J. Plant species endemism in savanna vegetation on table mountains (*Campo Rupestre*) in Brazil. *Vegetatio* **113**, 125–139 (1994).
43. Antonelli, A., Verola, C. F., Parisod, C. N. & Gustafsson, aL. S. Climate cooling promoted the expansion and radiation of a threatened group of South American orchids (Epidendroideae: Laeliinae). *Biol. J. Linn. Soc.* **100**, 597–607 (2010).
44. Freeman, D. C., Harper, K. T. & Charnov, E. L. Sex change in plants: old and new observations and new hypotheses. *Oecologia* **47**, 222–232 (1980).
45. Zimmerman, J. K. Ecological correlates of labile sex expression in the orchid *Catsetum viridiflavum*. *Ecology* **72**, 597–608 (1991).
46. Arditti, J. & Ghani, A. K. A. Numerical and physical properties of orchid seeds and their biological implications. *New Phytol.* **145**, 367–421 (2000).
47. Yukawa, T., Ogura-Tsujita, Y., Shefferson, R. P. & Yokoyama, J. Mycorrhizal diversity in *Apostasia* (Orchidaceae) indicates the origin and evolution of orchid mycorrhiza. *Am. J. Bot.* **96**, 1997–2009 (2009).
48. Gentry, A. H. & Dodson, C. H. Diversity and biogeography of Neotropical vascular epiphytes. *Ann. Missouri Bot. Gard.* **74**, 205–233 (1987).
49. Zotz, G. The systematic distribution of vascular epiphytes – a critical update. *Bot. J. Linn. Soc.* **171**, 453–481 (2013).

50. Chomicki, G. *et al.* The velamen protects photosynthetic orchid roots against UV-B damage, and a large dated phylogeny implies multiple gains and losses of this function during the Cenozoic. *New Phytol.* **205**, 1330–1341 (2015).
51. Salazar, G. A., Cabrera, L. I., Gerlach, G., Hågsater, E. & Chase, M. W. Phylogenetic relationships in *Mormodes* (Orchidaceae, Cymbidieae, Catasetinae) inferred from nuclear and plastid DNA sequences and morphology. *Phytotaxa* **263**, 18–30 (2016).
52. Monteiro, S. H. N., Selbach-Schnadelbach, A., de Oliveira, R. P. & van den Berg, C. Molecular phylogenetics of *Galeandra* (Orchidaceae: Catasetinae) based on plastid and nuclear DNA sequences. *Syst. Bot.* **35**, 476–486 (2010).
53. Pérez-Escobar, O. A., Parra-Sánchez, E. & Valdivieso, P. O. Inventario orquideológico de la Reserva Bosque de Yotoco, Valle del Cauca. *Acta Agronómica* **58**, 189–196 (2009).
54. Kolanowska, M., Pérez-Escobar, O. A., Parra-Sánchez, E. & Szlachetko, D. L. *An illustrated field guide to the orchids of the Yotoco Forest Reserve (Colombia)*. (Wydawnictwo Uniwersytetu Gdańskiego, 2011).
55. Batista, J. N. *et al.* Molecular phylogenetics of Neotropical *Cyanaeorchis* (Cymbidieae, Epidendroideae, Orchidaceae): geographical rather than morphological similarities plus a new specie. *Phytotaxa* **156**, 251–272 (2014).
56. Maddison, W. P. & FitzJohn, R. G. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Syst. Biol.* **64**, 127–136 (2015).
57. Davis, M. P., Midford, P. E. & Maddison, W. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evol. Biol.* **13**, 38–49 (2013).
58. Gregg, K. B. Variation in floral fragrances and morphology: incipient speciation in *Cycnoches*? *Bot. Gaz.* **144**, 566–576 (1983).
59. Parra-Sánchez, E., Retana, J. & Armenteras, D. Edge Influence on diversity of orchids in Andean cloud forests. *Forests* **7**, 1–13 (2016).
60. Carr, G. F. J. The genus *Cycnoches*: species and hybrids. *Orchid Rev.* 1–31 (2012).
61. Zachos, J. C., Dickens, G. R. & Zeebe, R. E. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**, 279–283 (2008).
62. Neubig, K. M. *et al.* Generic recircumscriptions of Oncidiinae (Orchidaceae: Cymbidieae) based on Maximum Likelihood analysis of combined DNA datasets. *Bot. J. Linn. Soc.* 117–146 (2012).
63. Bone, R. E., Smith, J. A. C., Arrigo, N. & Buerki, S. A macro-ecological perspective on crassulacean acid metabolism (CAM) photosynthesis evolution in Afro-Madagascan drylands: Eulophiinae orchids as a case study. *New Phytol.* **208**, 469–481 (2015).
64. Pérez-Escobar, O. A. *et al.* Andean mountain building did not preclude dispersal of lowland epiphytic orchids in the Neotropics. *Sci. Rep.* **7**, 4919 (2017).
65. Irimia, R.-E., A. Pérez-Escobar, O. & Gottschling, M. Strong biogeographic signal in the phylogenetic relationships of *Rochefortia* Sw. (Ehretiaceae, Boraginales). *Plant Syst. Evol.* **301**, 1509–1516 (2014).
66. Bechteler, J. *et al.* Geographical structure, narrow species ranges, and Cenozoic diversification in a pantropical clade of epiphyllous leafy liverworts. *Ecol. Evol.* **7**, 638–653 (2017).
67. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
68. Balbuena, J. A., Míguez-Lozano, R. & Blasco-Costa, I. PACo: a novel procrustes application to cophylogenetic analysis. *PLoS One* **8**, e61048 (2013).
69. Feldberg, K. *et al.* A phylogeny of Cephaloziaceae (Jungermanniopsida) based on nuclear and chloroplast DNA markers. *Org. Divers. Evol.* **16**, 727–742 (2016).
70. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772–772 (2012).
71. Stamatakis, A. RAXML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
72. Ronquist, F. *et al.* MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
73. Miller, M. A. *et al.* A RESTful API for access to phylogenetic tools via the CIPRES Science Gateway. *Evol. Bioinforma.* **11**, 43–48 (2015).
74. Erixon, P., Svennblad, B., Britton, T. & Oxelman, B. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Syst. Biol.* **52**, 665–673 (2003).
75. Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F. & Douzery, E. J. P. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol. Biol. Evol.* **20**, 248–254 (2003).
76. Drummond, A. J. & Bouckaert, R. R. *Bayesian evolutionary analysis with BEAST 2*. (Cambridge University Press, 2014).
77. Bouckaert, R. *et al.* BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10**, 1–6 (2014).
78. Ramírez, S. R., Gravendeel, B., Singer, R. B., Marshall, C. R. & Pierce, N. E. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* **448**, 1042–1045 (2007).
79. Poinar, G. Orchid pollinaria (Orchidaceae) attached to stingless bees (Hymenoptera: Apidae) in Dominican amber. *Neues Jahrb. für Geol. und Paläontologie* **297**, 287–293 (2016).
80. Töpel, M. *et al.* SpeciesGeoCoder: Fast categorisation of species occurrences for analyses of biodiversity, biogeography, ecology and evolution. *BioRxiv* **9274**, 1–15 (2014).
81. Givnish, T. J. *et al.* Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: insights from an eight-locus plastid phylogeny. *Am. J. Bot.* **98**, 872–895 (2011).
82. Matzke, N. J. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* **63**, 951–970 (2014).
83. Chamberlain, S., Ram, K. & Hart, T. spocc: interface to species occurrence data sources. *R Packag. version 0.4.5*. <http://CRAN.R-project.org/package=spocc> (2016).
84. Hijmans, R. J. & Elith, J. Species distribution modeling with R. (2016).
85. Arrigo, N. *et al.* A total evidence approach to understanding phylogenetic relationships and ecological diversity in *Selaginella* subg. *Tetragonostachys*. *Am. J. Bot.* **100**, 1672–1682 (2013).
86. Oksanen, J. *et al.* Package ‘vegan’. *R Packag. ver. 2*, 0–8, <https://doi.org/10.4135/9781412971874.n145> (2013).
87. Romero-González, G. A. & Nelson, C. E. Sexual dimorphism in *Catasetum* orchids: forcible pollen emplacement and male flower competition. *Science* **232**, 1538–1540 (1986).
88. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
89. Wickham, H. *GGPLOT2: elegant graphics for data analysis*. (Springer-Verlag, 2009). <https://doi.org/10.1007/978-0-387-98141-3>.
90. Gerlach, G. & Pérez-Escobar, O. A. Looking for missins swans: phylogenetics of *Cycnoches*. *Orchids* **83**, 434–437 (2014).
91. FitzJohn, R. G. Analysing diversification with diversitree. *R Packag. ver. 0*, 9–2 (2010).
92. Rabosky, D. L. & Goldberg, E. E. Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst. Biol.* **64**, 340–355 (2015).

Acknowledgements

We are grateful to Martina Silber (Munich, Germany) for valuable assistance provided during lab work, to Mario Blanco (San José, Costa Rica), Diego Bogarín (San José, Costa Rica), Norman Cash and family (Managua, Nicaragua) for assistance during field trips, and to Gustavo Romero (Herbarium HUH, Cambridge, U.S.A)

and Mario Blanco (University of Costa Rica) for providing plant material. Adarilda Petini-Bennelli provided vegetal material of Catasetinae. Paula Rudall and Richard Bateman critically read the manuscript. To R. Camara, for discussions. To the handling editor and two anonymous reviewers for providing critical comments to the manuscript. This work was supported by the Colombian Administrative Department of Science and Technology (COLCIENCIAS; granted to O.A.P.E., ref. 512), and the Deutsche Forschungsgemeinschaft (granted to GG; GE 828/12-1). F.L.C. has benefited from an “Investissements d’Avenir” grant managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01).

Author Contributions

O.A.P.E., G.C., F.L.C., J.M.dV. and J.H. designed research; A.M. and E.S. collected samples; A.M., and E.S. performed the lab work; O.A.P.E., F.L.C., and G.C. performed all analyses; O.A.P.E., G.C., F.L.C., J.M.dV., B.K., J.H., G.G., A.M., and E.S. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12300-y>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017