1

Multiple Geographical Origins of Environmental Sex Determination enhanced the

diversification of Darwin's Favourite Orchids

4

Oscar Alejandro Pérez-Escobar[1*]; Guillaume Chomicki[2]; Fabien L. Condamine[3];

Jurriaan M. de Vos[4*]; Aline C. Martins[5]; Eric C. Smidt[5]; Bente Klitgård[1]; Günter

Gerlach[6]; Jochen Heinrichs[2]

8

[1] *Department of Identification and Naming, Royal Botanic Gardens Kew,*

Richmond, TW9 3AB, UK

[2] *Department für Biologie I, Systematische Botanik und Mykologie, Ludwig-*

Maximilians-Universität, Menzinger Straße 67, D - 80638 München, Germany

[3] *CNRS, UMR 5554 Institut de Sciences de l'Evolution (Université de*

Montpellier), Place Eugène Bataillon, 34095 Montpellier, France

[4] *Department of Comparative Plant and Fungal Biology, Royal Botanic Gardens*

Kew, Richmond, TW9 3AB, United Kingdom

[5] *Department of Botany, Federal University of Paraná, PB 19031, Curitiba, PR*

81531-980, Brazil

[6] *Botanischer Garten München, Menzinger Straße 67, D - 80638 München,*

Germany

[*]*Corresponding author:* o.perez-escobar@kew.org; j.devos@kew.org

Running head: Origin and effect of sexual plasticity on Catasetinae diversification

Number of words: 3396 (main text) and 7052 (including references and captions)

Number of figures: 3

Number of tables: 1

Abstract [148 words]:

27   Environmental sex determination (ESD) – a change in sexual function during an

28   individual life span driven by environmental cues – is an exceedingly rare sexual

29   system among angiosperms. Because ESD can directly affect reproduction

30   success, it could influence diversification rate as compared with lineages that

31   have alternative mating systems. Here we test this hypothesis using a solid

32   phylogenetic framework of Neotropical Catasetinae, the angiosperm lineage

33   richest in taxa with ESD. We assess whether gains of ESD are associated with

34   higher diversification rates compared to lineages with alternative systems while

35   considering additional traits known to positively affect diversification rates in

36   orchids. We found that ESD has evolved asynchronously three times during the

37   last ~5 Myr. Lineages with ESD have consistently higher diversification rates

38   than related lineages with other sexual systems. Habitat fragmentation due to

39   mega-wetlands extinction, and climate instability are suggested as the driving

40   forces for ESD evolution.

41   **Key words:** Sexual plasticity, reproduction, historical biogeography,

42   Orchidaceae, diversification

43

44

45

46

47

48

49   **Introduction**

50   Angiosperms have evolved a rich array of sexual systems that are

51   unevenly distributed across lineages[1,2]. About 90% of angiosperms produce

52   male and female reproductive organs in the same flower[2–4]. A small fraction of

53   genera (i.e. 7%) include dioecious species, which bear sexes on separate

54   individuals[5]. One unusual sexual system is referred to as labile sex expression,

55   and involves changes in sex expression linked to environmental and phenotypic

56   interactions[6]. A regular form of labile sex expression is the recurrent alternation

57   of functionally male and female flower production over reproductive seasons[5]

58   (e.g. duodichogamy in *Acer*[6]). However, a more extreme form of labile sexual

59   expression is environmental sex determination (ESD), where an individual can

60   change its sex in response to environmental variables such as sunlight and

61   temperature[3,6,7]. Therefore, plants with ESD can produce male, female or

62   bisexual flowers either on the same individual or in separate individuals

63   depending on specific environmental conditions[8]. ESD is a rare sexual system in

64   angiosperms and has so far been reported for only *ca* 450 species in three

65   families[5]. Several authors proposed that habitat heterogeneity afforded by

66   fragmentation may provide conditions where ESD would be benefical[7–9]. Given

67   the differential demand of resources by male (e.g. pollen production) and female

68   (e.g. seed production) sex functions, ESD would enable the optimization of

69   resource reallocation to either sex function in response to specific site

70   characteristics.

71   Orchids, with about 25,000 species, are among the most species-rich

72   families of flowering plants on earth[10–12]. They are well known for their

73   diversity of reproductive systems[13], but these remain poorly understood[14,15]. The

74   vast majority of Orchidaceae produce male and female reproductive organs in

75   the same flower[16]. Indeed, orchid flowers exhibit an extreme case of

76   hermaphroditism in the plant kingdom because of the fusion of male and female

77   organs into a gynostemium[17,18]. An exception to this rule however are the

78   Catasetinae, a Neotropical orchid lineage containing 354 species[19,20]. Because of

79   their striking sexual systems (i.e. ESD, protandry, adichogamy) and the diversity

80   of pollination syndromes (e.g. male euglossine-bee pollination, oil bee

81   pollination)[21,22], Catasetinae have long attracted the attention of botanists and

82   naturalists[23]. However, we still know little about the origin and evolution of

83   ESD in this lineage.

84   Catasetinae includes eight genera, namely *Catasetum, Clowesia,*

85   *Cyanaeorchis, Cycnoches, Dressleria, Galeandra, Grobya* and *Mormodes*, and

86   most of the diversity is found in Central America and Amazonia regions[24]. Sex

87   expression in Catasetinae, as in several other plant lineages with ESD, is

88   regulated by ethylene production. The pioneer work of Gregg[25] documented the

89   role of sunlight exposure and ethylene synthesis on sex determination in

90   *Cycnoches* and *Catasetum.* She convincingly demonstrated that ethylene

91   concentrations are about two- to 100-fold higher in inflorescence apices that are

92   fully exposed to sunlight compared with those kept in shade. Consequently,

93   inflorescences with higher ethylene concentration usually produce female

94   flowers, whereas those with lower levels produce male flowers[25].

95   To date, about one third of the Catasetinae (i.e. *ca* 100 species) are

96   known to have ESD. Therefore, the group provides a unique opportunity to

97   study its evolutionary context in a phylogenetic framework. A recent study

98   provided strong support for three independent origins of sexual plasticity in

99   Catasetinae[14,26]. ESD evolved in *Catasetum, Cycnoches* and *Mormodes*, always

100   from a protandrous ancestor. Protandry, in turn, was found to be the ancestral

101   condition in the so-called core Catasetinae (*Catasetum, Clowesia, Cycnoches,*

102   *Dressleria* and *Mormodes*). ESD could be a switch in the evolution of dioecy

103   (i.e. spatial separation of sexes)[21], as observed in other plant lineages with sexual

104   plasticity such as *Fuchsia* and *Hebe*[27–29].

105   It is still unclear whether the independent origins of ESD occurred

106   synchronously and under the same environmental conditions in a single

107   biogeographical region or evolved asynchronously in different regions, possibly

108   under different environmental conditions. It also remains elusive whether

109   lineages with ESD are associated with higher diversification rates compared

110   with hermaphroditic lineages. In view of its rarity, we hypothesise that ESD

111   could be a disadvantageous system leading to lower diversification rates.

112   Alternatively, ESD could be evolutionarily unstable and could thus represent a

113   transitional state rather than a stable reproductive system.

114   The extremely high orchid diversity is thought to be related with the

115   evolution of novel traits and the invasion of novel environments[30], possibly

116   confounding analysis of the effects of ESD in orchid diversification. Therefore,

117   we also investigated the impact of epiphytism[10,31] and male euglossine-bee

118   pollination[21,30] on the diversification dynamics of Catasetinae. Epiphytism and

119   male euglossine-bee pollination are traits that occur in all ESD species of

120   Catasetinae, but these traits are shared with some Catasetinae species that have

121   alternative reproductive systems[24] (see *Methods*). Therefore, it is possible that

122   these traits might have influenced the diversification of the subtribe, yet their

123   specific contribution to diversification rates are still unknown.

124   The limited knowledge on the evolutionary dynamics of ESD in orchids

125   stems partly from the lack of well-sampled species-level phylogenies[5]. This is

126   particularly true for Catasetinae, in which previous studies did not attain even

127   taxon sampling, thus leaving species-rich clades dramatically underrepresented

128   (e.g. *Catasetum*; see[32–34]). In the current study, we generate the first densely

129   sampled phylogeny of Catasetinae. We combine data on species distribution and

130   a wide range of comparative phylogenetic methods to reconstruct the

131   evolutionary history of ESD. We then infer the biogeographic history of

132   Catasetinae, test whether their ESD lineages have higher diversification rates

133   than non-ESD lineages, while accounting for the possible contribution of other

134   traits to orchid diversification. We also investigate whether past and modern

135   climate niche preferences of lineages with ESD are divergent from non-ESD

136   lineages.

137

138   **Results**

139   *Phylogeny and molecular dating of* Catasetinae

140   Our DNA matrix included a total of 282 species, of which 132 belonged

141   to Catasetinae (i.e. ~37% of the extant diversity of the subtribe; Tab. S1).

142   Analyses based on Maximum Likelihood (ML) and Bayesian inference (BI)

143   recovered similar topologies. However, they revealed several conflicting

144   positions of species in the nuclear and plastid trees with bootstrap percentage

145   values (MLBS) $\geq$ 75 and/or posterior probabilities (PP) $\geq$ 0.95 (Fig. S1-S2).

146   Figure 1 and Figure S3 show the ML trees inferred from concatenated nuclear

147   and non-conflicting plastid sequences, respectively. Here, after exclusion of

148 conflicting positions from the DNA matrices (see *Methods* and Appendix S1 for

149 a detailed explanation on incongruence handling), almost all backbone nodes of

150 the phylogeny achieved MLBS $\geq$ 75 and PP > 0.95 support.

151 Divergence time estimates performed under a Bayesian relaxed clock

152 model yielded a mean Coefficient of Variation (CV) value for branch-specific

153 substitution rates of 0.73, indicating substantial among-branch rate

154 heterogeneity, confirming that the use of a relaxed molecular clock is

155 appropriate. Mean ages and the associated 95% credibility intervals obtained

156 under the relaxed clock model for all nodes are provided in Figure S4. The

157 dating analyses showed that *(i)* the Catasetinae and Eulophiinae+Cyrtopodiinae

158 shared a common ancestor in a time interval ranging from the early Miocene to

159 the early Oligocene (25.93 Ma $\pm$ 5 Myrs), and *(ii)* the diversification of the

160 Catasetinae most recent common ancestor (MRCA) started around 19.5 Ma $\pm$ 4

161 Myrs during the early Miocene.

162

163 *Geographical origin of ESD and biogeographical history of Catasetinae*

164 Table S4 provides ML statistics for six biogeographical models as

165 inferred using BIOGEOBEARS. The best fitting model was BayArea, including

166 founder-event speciation (*J* parameter; $LnL_{best\ model}$ = - 860.3 vs $LnL_{second\ best}$

167 $_{model}$= -875.5; $AIC_{best\ model}$ = 1727 vs $AIC_{second\ best\ model}$ = 1757; Tab. S4). This

168 model estimated south-eastern South America as the most likely ancestral area

169 of Catasetinae (Fig. 2; Fig. S5). The same region was supported as the ancestral

170 area of *Cyanaeorchis* and *Grobya*. In contrast, *Galeandra* and *Catasetum*

171 inhabited larger areas between Amazonia+Guiana Shield+south-eastern South

172   America and Amazonia+south-eastern South America, respectively. The origins

173   of *Clowesia* and *Cycnoches* were estimated to be in Amazonia, whereas the

174   remaining lineages of Catasetinae (i.e. *Dressleria*, *Mormodes*) originated in

175   Central America (Fig. 2; Fig. S5).

176

177   *Origin, climate niche and evolutionary dynamics of ESD*

178   Ancestral character estimation analysis under the best model (i.e. single

179   transition rate) identified three independent gains of environmental sex

180   determination, without any losses (Fig. 2; Fig. S6). ESD first evolved during the

181   early Pleistocene to late Miocene (5 Ma ±3 Myrs) at the MRCA of *Cycnoches* in

182   the Amazonia region (Fig. 2; Fig. S5). Subsequently, two gains of ESD occurred

183   towards the Pleistocene to Pliocene (3 Ma ±1 Myrs) and the Pleistocene (1 Ma

184   ±1 Myrs) in Amazonia – south-eastern South America and Central America

185   regions, respectively. These gains took place at the MRCAs of *Catasetum* and a

186   small clade of *Mormodes* species restricted to Central America and Northern

187   Andes, respectively (Fig. 2; Fig. S5, S6). Lineages with ESD further diversified

188   mostly in lowland areas (i.e. Amazonia, Choco, south-eastern South America,

189   Central America, Guiana Shield. Two episodic colonisations from Central

190   America towards Northern Andes were recovered.

191   Bioclimatic variables for 1372 georeferenced occurrences of Catasetinae

192   were adequately captured by the first two axes of a principal component analysis

193   (PCA) (47.9% and 23.6%, respectively). The first PCA axis received

194   contributions from all bioclimatic variables except BIO 3 (isothermality), BIO 4

195   (temperature seasonality), BIO 7 (temperature annual range), BIO 15

196   (precipitation seasonality), and altitude (Fig. S7). The second PCA axis mostly

197   discriminated dry/hot environments, but also included contributions from BIO

198   15 (precipitation seasonality; Fig. S8). Therefore, discriminations between

199   particular environments were not detected (Fig. S7-S9), as bioclimatic variables

200   representing precipitation and temperature contributed similar amounts of

201   variance to PCA axes 1 and 2 (**Fig. S7, S8**). Overall, the ordination of climatic

202   niche space of lineages based on the bioclimatic variables indicated no evidence

203   for general ecological segregation between lineages with and without ESD (**Fig.**

204   **S9a**).

205   To further assess the apparent lack of climatic niche differentiation

206   between ESD and non-ESD taxa, we excluded all auto-correlated variables (see

207   *Methods*), and performed a non-metric dimensional analyses (NMDS) based on

208   altitude, BIO 15 (precipitation seasonality), BIO 18 (precipitation of warmest

209   quarter) and BIO 19 (precipitation of coldest quarter). Similarly, we found

210   extensive overlaps between climatic niches of ESD and non-ESD lineages (Fig.

211   S9b). However, climatic niche differentiation analysis using present-day

212   bioclimatic variables are not directly informative of the climate niche

213   preferences of ancestral lineages in which ESD arose. Thus, we performed

214   ancestral character estimation analyses of twenty bioclimatic variables to assess

215   whether ancestors having evolved ESD and those with alternative sexual

216   systems diverged in their climatic niche preferences. We found that ancestral

217   mean values of all bioclimatic variables were consistently similar between

218   ancestors that independently evolved ESD and contemporary lineages with other

219   sexual systems (Fig. 3). More importantly, ancestral mean bioclimatic values of

220   lineages with and without ESD revealed little variation through time (Fig. S10).

221   Trait-dependent diversification rate analysis supported a model with free

222   speciation rates in lineages with and without ESD ($LnL_{best\ model}$=-248.85;

223    $AICc_{best\ model}$=506.035 vs. $LnL_{null\ model}$=-262.41; $AICc_{null\ model}$=531.022, Tab. 1;

224   Tab. S5). Under the best model, speciation rates were ~1.6-fold higher in

225   lineages with ESD, while extinction rates were modelled under the same

226   parameter (Fig. S11a; Tab. 1). We confirmed the robustness of these results

227   through our tailored data simulations (see *Methods*), that generated an

228   expectation about the distribution of ΔAIC values obtained from randomly

229   reshuffled trait data (see *Extended Materials and Methods* of Appendix S1). The

230   distribution of simulated ΔAIC did not overlap empirical AIC values. This

231   suggests that our trait-dependent analysis is not biased by type I error (Fig.

232   S14a).

233

234   *Evolutionary dynamics of male euglossine-bee pollination and epiphytism*

235   Binary-state dependent diversification analyses of male euglossine-bee

236   pollination and epiphytism supported models with free speciation rates and

237   equal extinction and transition rates (M_2; see *Methods*), and free speciation and

238   transition rates but equal extinction rates (M_6), respectively. Both models

239   recovered slightly higher speciation rates for male euglossine-bee pollination

240   (2.13 events/Myr; Fig. S12; Tab. 1; Tab. S6), and epiphytism (1.56 events/Myr;

241   Fig. S13; Tab. 1; Tab. S7) than for lineages with alternative reproductive

242   systems (1.82 events/Myr; Tab. S6) and plant habit (1.27 events/Myr; Tab. S7).

243   Analysis of simulated data confirmed the robustness of real data analysis results

244   for male euglossine-bee pollination. The distribution of ΔAIC values from real

245   data and simulated data did not overlap: $ΔAIC_{simulated}$ was centred towards 30

246   units while the $\Delta AIC_{real\ data}$ was estimated at 9 (Fig. S14b). Analysis of simulated

247   data for epiphytism did not reach convergence, hence the veracity of binary-state

248   dependent diversification analysis on the real data set could not be evaluated.

249

250   **Discussion**

251   *Forest fragmentation as the context for the multiple geographical origins of*

252   *ESD*

253   Orchids have a rich variety of reproductive systems[35], of which

254   hermaphroditism is the most widespread in the family[16,35], reflecting the

255   angiosperm-wide pattern[2]. The Catasetinae, however, have evolved three

256   reproductive systems (i.e. adichogamy, dichogamy, and ESD[14]), thus providing

257   a unique opportunity to investigate biotic and abiotic factors shaping the

258   evolution of these sexual systems. By relying on a solid phylogenetic

259   framework, divergence time estimations, and ancestral character estimation, our

260   study demonstrates the independent temporal and geographic origins of ESD in

261   Catasetinae. More importantly, it reveals for the first time an apparent positive

262   correlation between gains of sexual plasticity and higher speciation rates.

263   However, our analyses show that other traits (i.e. male euglossine-bee

264   pollination[21,30], and perhaps epiphytism[30]), may also have positively contributed

265   to the diversification of the Catasetinae.

266   Ancestral area estimation analyses suggest that sexual plasticity first

267   originated in the Amazonian region during the late Miocene (5 Ma ±3 Myrs), a

268   period preceded by the disappearance of large aquatic systems that dominated

269   Amazonia from 17 to 7 Ma[36]. Such aquatic environments, known as Pebas Lake

270   and Acre Systems, ranged from the drainage of western Amazonia to the

271   Caribbean coast of Venezuela[37,38] and prompted the fragmentation of pre-

272   existing forests. The disappearance of Pebas and subsequent aquatic systems

273   towards the late Miocene gave rise to flood-plains with a patchy vegetation

274   dominated by palms, ferns and grasses[38,39].

275   The late Miocene vegetation changes in the Amazonian region may also

276   be related to the global cooling subsequent to the mid-Miocene Climatic

277   Optimum (~15 Ma). These climate changes might have promoted forest

278   fragmentation by range expansions and contractions of species due to dramatic

279   temperature changes[37,40,41]. An example of the effect of drastic temperature

280   changes on vegetation ranges is the peculiar *Campos Rupestres* in Brazil. There,

281   dry periods provided the optimal conditions for vegetation to expand, allowing

282   species to expand their distribution ranges, and subsequently occupy larger

283   areas. Succeeding wetter periods would have prompted the contraction of

284   distribution ranges, leading to the formation of isolated vegetation patches[42,43].

285   As a result, species like the orchids *Sarcoglottis caudata* and *Liparis beckeri*

286   now occupy very restricted ranges, whereas other taxa have a disjunctive

287   distribution, e.g. *Vellozia virgata.*

288   Sexual plasticity has been regarded as an adaption to patchy

289   environments[7,8,44], and this might have explanatory power for the gains of sexual

290   plasticity in Catasetinae. The existence of patchy forests towards the end of the

291   late Miocene in South America and vegetation changes in Central America

292   might have facilitated the evolution of ESD by leading individuals to invest

293   different amounts of resources to male and female gamete production. Resource

294   reallocation to either function might thus have taken place as a function of an

295   individual's specific site characteristics[9,44].

296   In a seminal paper, Zimmerman[45] provided evidence towards the biased

297   sex expression in *Catasetum viridiflavum.* In this species, female plants are more

298   frequently distributed in open canopies whereas male plants often grow in closed

299   canopies. Plants growing in open canopies are longer exposed to sunlight

300   irradiation, thus favouring accumulation of energetic resources (i.e. plant

301   biomass) that are later reallocated to female flower and fruit production[45]. The

302   latter process requires considerable amounts of energy[46] because of the massive

303   number of seeds produced in a single capsule (see below). Sunlight irradiation

304   appears to be the main ecological variable driving male/female flower

305   production ratio in Catasetinae[25,45]. However, how other variables tied to

306   resource constrains and patch quality (e.g. substrate type[45], orchid-mychorriza

307   interactions[47]) affect sex ratios in the Catasetinae deserves further research.

308   *ESD, epiphytism and male euglossine-bee pollination enhanced*

309   *diversification in Catasetinae*

310   The overwhelmingly low frequency of lineages with ESD compared to

311   the vast majority of hermaphrodite angiosperms[5] could suggest that sexual

312   plasticity is a disadvantageous mating system or that it is evolutionary unstable.

313   However, our study shows that Catasetinae clades with ESD have higher

314   speciation rates than those with other sexual systems (i.e. adichogamy,

315   protandry; Fig. S11). Here, the overarching comparison of speciation rates

316   obtained from all analysed traits revealed a greater contribution of ESD

317   evolution to the diversification dynamics of Catasetinae (rates ~1.6-fold higher

318   than in lineages with other sexual systems). In contrast, the evolution of

319   epiphytism (~1.2-fold higher) and male euglossine-bee pollination (~1.17-fold

320   higher) affected diversification in the Catasetinae, but to a lesser extent.

321   Our ancestral character estimation analysis does not reveal any loss of

322   ESD, suggesting that it is an evolutionary stable system in Catasetinae that has

323   been maintained for up to 5 million years. Nevertheless, diversification in

324   Catasetinae has also been prompted by the gain of traits that were earlier shown

325   to influence diversification rates of Neotropical orchids in general (i.e. male

326   euglossine-bee pollination[21], epiphytism[30,31]). Taken together, these findings

327   indicate that rate clade diversification is determined by several, potentially

328   interacting factors[30].

329   Vascular epiphytes are a key component of tropical forests[48], and orchids

330   account for 68% of the total vascular epiphyte diversity[49]. Epiphytism has

331   evolved at least seven times since the Eocene in Orchidaceae[50], and might have

332   enhanced diversification by allowing colonization of new, largely unoccupied

333   habitats such as branches and trunks of angiosperms[30]. In Catasetinae,

334   epiphytism occurs in all members of the three most speciose genera, namely

335   *Catasetum*, *Cycnoches* and *Mormodes*[24,51] (290 species altogether[20]). It also

336   occurs in all species of *Grobya*[24] (five species), *Clowesia* (seven species) and six

337   of the eleven species of *Galeandra*[52]. Epiphytic species from such clades inhabit

338   in a wide array of ecosystems and biogeographical regions[24,53,54]. Contrastingly,

339   terrestrial species are confined to species-poor clades with restricted distribution

340   such as *Dressleria* (twelve species), *Cyanaeorchis* (three species), and a clade of

341   five species nested in *Galeandra*[24,52,55].

342   Pollination by male euglossine-bee evolved independently three times in

343   Orchidaceae since the early Miocene[21]. It may have accelerated diversification

344   rates by promoting switches in pollinators via small changes in the chemical

345   profile of the floral scent, or alternatively by changes in pollinarium placement

346   sites, hence allowing reproductive isolation[30]. Such pollination syndrome occurs

347   in all species of *Catasetum, Clowesia, Cycnoches, Dressleria* and *Mormodes*,

348   which altogether make up ~85% of total species number in Catasetinae.

349   Alternative pollination syndromes (e.g. oil-bee pollination[22], food deception)

350   occur in the remaining, species-poor genera of Catasetinae (i.e. *Cyanaeorchis,*

351   *Galeandra, Grobya*). Taken together, our results point to a prominent role of

352   ESD in the evolutionary dynamics of the Catasetinae, evidenced by speciation

353   rates ~1.6-fold higher than in orchids without ESD. Nonetheless, we do not

354   exclude the additional contributions of traits, such as euglossine-bee pollination

355   and epiphytism.

356   The reliability of our diversification rate results is strongly dependent on

357   the power and hypotheses testing ability of our trait-dependant diversification

358   model (i.e. BiSSE). The BiSSE method is known to be affected by within-clade

359   pseudoreplication[56], and sample size[57], issues that are inherent to the number of

360   independent origins of the analysed traits, and the sampling proportion of the

361   clade under study, respectively. Thus, extreme care must be used when

362   interpreting the outcome of trait-dependant diversification analyses[57]. However,

363   by comparing different trait-diversification models with randomly reshuffled

364   coding matrices in our simulations, and analysing other traits known to affect

365   diversification rates in orchids, we have partially accounted for pseudo-

366   replication bias. Indeed, our customized simulations revealed that BiSSE

367   analyses of ESD and male euglossine-bee pollination are not affected by Type I

368   errors. Nonetheless, diversification rate analyses can only account for the effects

369   of the traits here considered, hence we cannot entirely rule out that other

370   variables have additional effects[56].

371

372   *ESD, protandry and adichogamy share the same climatic niche in Catasetinae*

373   We have not detected significant differences in the ancestral and modern

374   climate niches between Catasetinae lineages with ESD and other sexual systems

375   (Fig. 3; Figs. S9-S10). Thus, particular temperature or precipitation regimes do

376   not appear to have favoured any of these lineages. Species with ESD that occur

377   often in habitats with longer sunlight exposure compared to species without

378   ESD, produce bigger pseudobulbs, female flowers and fruits[45,58]. These species

379   also produce larger capsules than Catasetinae lineages with protandry or

380   adichogamy (e.g. *Cyanaeorchis, Clowesia, Dressleria*; Pérez-Escobar, pers.

381   obs.). For instance, the capsules of *Cycnoches chlorochilon* produce on average

382   up to three times more seeds (3,770,000) than fruits of *Cymbidium tracyanum*

383   (850,000), an adichogamous species closely related to Catasetinae[46]. Therefore,

384   under the same climatic niche, Catasetinae lineages with ESD might have

385   produced more offspring than lineages with other sexual systems. However,

386   additional data on fruit size and particularly on seed content for all Catasetinae

387   members are needed to statistically test such assumption.

388   Occurrence of ESD requires a heterogeneous habitat (e.g. fragmented

389   forest matrix), because under homogenous ecological conditions (e.g. forest

390   understory) plants with ESD will bear predominantly one sex form, and mating

391   success may thus be dramatically reduced. This highlights the important role of

392   climate instability and its effects on habitat heterogeneity since the Miocene

393   Climate Optimum in the evolution and maintenance of ESD. Nevertheless, it is

394   likely that bioclimatic variables do not represent relevant environmental aspects

395   of the niche in which ESD orchids occurs as compared with non-ESD lineages.

396   Also, the climatic niche analyses pinpoint the explanatory power of variables

397   with maximum variance, hence they will be mostly represented in PCA axes 1

398   and 2. However, such variables are not necessarily the most relevant and

399   informative for our group of study. Therefore, even though ESD and non-ESD

400   species niches largely overlap, ESD lineages might still be dependent on more

401   specific ecological variables that are not necessarily captured by bioclimatic

402   variables but that may yet be driven by climate (e.g. those operating on forest

403   edge matrices[59], or affecting physiognomic properties of vegetation).

404

405   **Conclusions**

406   Environmental Sex Determination (ESD) in Catasetinae has long

407   attracted and even confused[60] botanists and naturalists[23], and the evolutionary

408   and biogeographic context for the evolution of this remarkable trait remained

409   elusive before our study. Our results show that the evolution of ESD involved

410   multiple gains in three biogeographical regions between late Miocene and late

411   Pleistocene - geological periods that were marked by the extinction of pivotal

412   aquatic environments and mega-wetlands[36,39] and dramatic fluctuations in

413   temperature[61]. Fragmentation of vegetation prompted by the demise of these

414   aquatic systems in South America and by dramatic changes in temperature

415   across the Neotropics might have fostered the evolution of sexual plasticity in

416   Catasetinae. In spite of the extreme rarity of sexual plasticity among the rich

417   array of angiosperm reproductive systems[5], evolution of ESD in Catasetinae

418   appears to be correlated with higher speciation rates. Other traits such as male

419   euglossine-bee pollination and epiphytism, however, also appear to have

420   positively influenced the spectacular diversity of Catasetinae. This illustrates the

421   difficulty in isolating the effect of one trait on diversification rates from effects

422   of others. The genetic contribution of an individual to future generations is

423   enhanced by ESD because it promotes the development of the sex most suited

424   for specific environmental conditions (i.e. sunlight intensity in Catasetinae[8,44]).

425   We demonstrate that ESD-bearing Catasetinae inhabit the same climatic niche as

426   their non-ESD relatives. Thus, plants with ESD might produce more offspring

427   over plants having alternative sexual systems because of the increased

428   opportunities for optimizing resource allocation to male and female function at

429   each growing site.

430

431   **Materials and methods**

432   *Taxon sampling, DNA sequencing and phylogenetic analyses*

433   Species names, geographical origins, voucher specimens, and GenBank

434   accession numbers of sequences included in phylogenetic analyses are provided

435   in Table S1. We built our alignment based on the molecular datasets of Whitten

436   et al.[33], Salazar et al.[51], Neubig et al.[62], Bone et al.[63], and Pérez-Escobar et

437   al.[34,64], and further generated 57 nuclear ribosomal and 25 plastid sequences for

438   poorly sampled lineages such as *Catasetum, Dressleria* and *Grobya* (Tab. S1).

439   Because the precise position of Catasetinae inside the tribe Cymbidieae was

440   elusive[33,55], we also sampled selected representatives of closely related subtribes

441   (i.e. Cyrtopodiinae, Dipodiinae, Eulophiinae, and Oncidiinae). Genomic DNA

442   extraction, amplification, sequencing and alignment procedures are the same as

443   in Irimia et al.[65], Pérez-Escobar et al.[14], and Bechteler et al.[66]. For the newly

444   sampled taxa, we sequenced nuclear ribosomal external and internal transcribed

445   spacers (ETS and ITS, respectively), a fragment of the *Xdh* gene, and also a

446   ~1500 bp fragment of the plastid gene *ycf*1, and the *trn*S–*trn*G intergenic spacer.

447   Amplification settings and sequencing primers used for ITS, ETS, *Xdh*, *trn*S–

448   *trn*G, and *ycf*1 are specified in Table S2.

449   Loci were aligned using MAFFT 7.1[67] and were further manually checked,

450   yielding a total matrix of 281 taxa and 8104 nucleotides (~24% parsimony informative

451   sites). Congruence between nuclear and plastid datasets was assessed following Pérez-

452   Escobar et al.[34], and using PACo[68] (Appendix S1). The procedure is available as a

453   pipeline (http://www.uv.es/cophylpaco/) and was also employed to identify operational

454   terminal units (OTUs) from the plastid dataset conflicting with the nuclear dataset

455   (potential outliers detected by PACo are shown in Fig. S1-S2). A detailed explanation

456   on PACo and a rationale on outlier handling is provided in the Extended Materials and

457   Methods section of Appendix S1.

458   Phylogenetic analyses of separate and concatenated loci were carried out

459   under maximum likelihood (ML) and Bayesian inference (BI), following Pérez-

460   Escobar et al.[14,34], and Feldberg et al.[69]. The best-fitting substitution models for

461   ML and Bayesian analyses were obtained for each data partition using

462   jModelTest v.2.1.6[70] (Tab. S3). Phylogenetic inference relied on the ML and BI

463   approaches implemented in RAxML-HPC 8.2.4[71] and MrBayes 3.2.2[72],

464   respectively, and were carried out on the CIPRES Science Gateway computing

465   facility[73]. Maximum likelihood bootstrap supports (MLBS) were generated for

466   the ML tree and MLBS ≥ 75% considered as good support, and Bayesian

467   Posterior Probabilities (PP) ≥ 0.95 for the Bayesian majority rule consensus

468   topology were regarded as significant[74,75].

469

470   *Molecular clock dating*

471   Divergence time estimates were conducted using the Bayesian relaxed

472   clock approach of BEAST 2.1.3[76] with a concatenated nuclear+plastid subset of

473   the data obtained after the PACo analysis. Strict and uncorrelated lognormal

474   molecular clock models, both with pure-birth speciation models as

475   recommended for species level sampling[77], were compared to explore clock-

476   likeness of the data. For calibrating the relaxed clock model, there are few

477   fossils unambiguously assigned to Orchidaceae[78], and these are placed to

478   lineages very distantly related to Catasetinae (i.e. *Dendrobium*, *Earina*, both

479   Vandeae). Recently, a new putative orchid fossil from the Dominican amber

480   (early - middle Miocene, 15-20 Ma) was described and assigned to the tribe

481   Cymbidieae[79]. However, the author provided no evidence to unambiguously

482   assign such fossil to Cymbidieae or orchids in general. Secondary calibrations

483   are therefore the only option for dating the tree of Catasetinae. To this end, we

484   relied on the age estimates obtained from a fossil-calibrated global phylogeny of

485   Orchidaceae[50], which were set with normally distributed priors that reflected

486   uncertainty in the primary analysis[76]. Settings for absolute age estimation are

487   detailed in the *Extended Materials and Methods* of Appendix S1.

488

489   *Ancestral areas estimation*

490   Species ranges were coded from the literature[19,60] and from herbarium

491   specimens housed in the herbaria AMES, COL, F, M, MO, SEL, US, using the

492   R-package SpeciesGeoCoder[80]. Distribution data were also obtained from own

493   field observations. Biogeographical areas were defined considering putative

494   areas of endemism of Catasetinae as well as species distributions observed in

495   other plant lineages such as Rubiaceae[37] and Bromeliaceae[81]. We divided the

496   geographical range of Catasetinae and outgroup taxa in eight biogeographic

497   regions (see *Extended Materials and Methods* of Appendix S1 for a detailed

498   description of coded areas): (1) Central America; (2) Guiana Shield; (3)

499   Amazonia[37]; (4) Chocó; (5) Northern Andes; (6) Central Andes; (7) South-

500   eastern South America; and (8) Africa and Australasia. We used the R-package

501   BIOGEOBEARS (Biogeography with Bayesian and Likelihood Evolutionary

502   Analysis in R script[82]) to estimate ancestral areas on the phylogeny of

503   Catasetinae.

504

505   *Climatic niche analyses*

506   We mapped georeferenced collection records obtained from floras, GBIF

507   database and herbarium specimens (mean of five specimens per species,

508   maximum number of record per species set to ten), and they represent the known

509   distribution of extant Catasetinae species included in our taxon sampling. To

510   query GBIF database, we relied on the function *occ* of the R-package SPOCC[83].

511   We further extracted corresponding values of altitude and 19 climatic variables

512   (30 seconds resolution) reflecting temperature and precipitation regimes from

513   the WorldClim database (available at: http://www.worldclim.org/current), using

514   the function *extract* of the R-package RASTER[84]. To characterise the climate-

515   niche occupied by Catasetinae lineages with ESD or other sexual systems, we

516   performed Principal Component Analysis (PCA) on the mean values of the 20

517   environmental variables of each species, using the default function *princomp* of

518   the R (Fig. S7). We further explored the contribution of bioclimatic variables to

519   axes with the most variance by plotting the loadings of every bioclimatic

520   variable onto the axes (Fig. S8, S9).

521   To avoid spurious results arising from inclusion of correlated variables

522   [63,85], we determined the Pearson's correlation coefficients between the variables

523   and altitude and then included only variables with a Pearson's correlation

524   coefficient <0.5, taking a single variable in correlated clusters. This way, we

525   selected altitude, BIO 15 (precipitation seasonality), BIO 18 (precipitation of

526   warmest quarter) and BIO 19 (precipitation of coldest quarter) as a set of

527   maximally uncorrelated variables. We analysed these variables using the R-

528   package VEGAN[86] to perform non-dimensional metric scaling analyses

529   (NMDS) using the dataset of 1372 georeferenced herbarium specimens.

530

531   *Ancestral character estimation*

532   We coded for absence (state 0) and presence (state 1) of ESD in 131

533   species of Catasetinae, plus eight taxa of Cyrtopodiinae (see below). We

534   sampled 68 of the 164 known species with ESD in Catasetinae (i.e. ~41% of the

535   ESD diversity). Information on occurrence of ESD was obtained from the

536   literature[24,58,87]. We estimated the origin of ESD within Catasetinae using a

537   stochastic character mapping approach implemented on a maximum likelihood

538   framework for ancestral character estimation analysis (ACE) implemented by

539   the function *make.simmap* in the R-package PHYTOOLS[88]. Under this

540   approach, we fitted single (ER), symmetrical (SYM), and asymmetrical

541   character transition rate models (ARD) and performed stochastic mapping on

542   1000 iterations using the maximum clade credibility tree derived from the

543   BEAST dating analysis (see above). Eight members of the Neotropical genus

544   *Cyrtopodium* were selected as outgroup taxa, and they exhibit hermaphroditic

545   flowers (i.e. no ESD).

546   To assess whether ancestors with ESD and those with alternative sexual

547   systems differed in their climatic niche preferences back in time, we performed

548   ACE analyses of the same nineteen bioclimatic variables plus altitude, obtained

549   from georeferenced occurrences (see *Climate niche analyses* section of

550   *Methods*). ACE of absolute mean values relied on a Maximum Likelihood

551   framework, and were performed on a phylogeny derived from the character

552   stochastic mapping analysis of ESD (using ER as transition rate model, see

553   *Results*). We then visualized the variation through time of all variables (**Fig. 3**),

554   using the function *phenogram* of the R-package PHYTOOLS[88]. We also plotted

555   mean ancestral values and their corresponding 95% confidence intervals (CI)

556   estimated for every node of the Catasetinae phylogeny, using the R-package

557   GGPLOT2[89] (Fig. S10).

558

559   *Trait-dependent diversification analyses*

560   We tested whether linages with ESD (state 1) had higher speciation rates

561   than hermaphroditic, monoecious clades (state 0). In addition, to tease apart the

562   contribution of biotic variables to Catasetinae diversification that are known to

563   positively influence diversification rates in orchids, we investigated the

564   evolutionary dynamics of male euglossine-bee pollination and epiphytism in the

565   subtribe. The former trait occurs in all species of the genera *Catasetum*,

566   *Clowesia*, *Cycnoches*, *Mormodes* and *Dressleria*[21,24,90], while the latter occur in

567   all species of *Catasetum*, *Clowesia*, *Cycnoches*, *Mormodes*, *Dressleria*[24], and

568   part of *Galeandra* species[52]. We relied on the Binary State Speciation and

569   Extinction (BiSSE) model[91] to estimate diversification rates associated with each

570   trait. We used the R-package DIVERSITREE[91] to perform eight diversification

571   models in the maximum likelihood framework on 100 randomly sampled dated

572   trees from the Bayesian divergence times analysis. We tested eight models with

573   different configurations of speciation, extinction, and transition rates between

574   characters. Detailed settings for binary-trait dependent diversification analyses,

575   and simulations to account for type I error biases[92] are provided in Appendix S1.

576   **References**

577   1.   Barrett, S. C. H. The evolution of plant reproductive systems: how often are
578   transitions irreversible? *Proc. R. Soc. B Biol. Sci.* **280,** 1–9 (2013).

579   2.   Barrett, S. C. H. The evolution of plant sexual diversity. *Nat. Rev. Genet.* **3,** 274–
580   284 (2002).

581   3.   Tanurdzic, M. & Banks, J. A. Sex-determining mechanisms in land plants. *Plant
582   Cell* **16,** 61–72 (2004).

583   4.   Bateman, R. M., Hilton, J. & Rudall, P. J. in *Flowers on the tree of life* (eds.
584   Wanntorp, L. & De Craene, L.) 315 (Cambridge University Press, 2011).

585   5.   Renner, S. S. The relative and absolute frequencies of angiosperm sexual
586   systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am. J.
587   Bot.* **101,** 1588–1596 (2014).

588   6.   Renner, S. S., Beenken, L., Grimm, G. W., Kocyan, A. & Ricklefs, R. E. The
589   evolution of dioecy, heterodichogamy, and labile sex expression in *Acer*.
590   *Evolution* **61,** 2701–2719 (2007).

591   7.   Charnov, E. L. & Bull, J. When is sex environmentally determined? *Nature* **266,**
592   828–830 (1977).

593   8.   Korpelainen, H. Labile sex expression in plants. *Biol. Rev. Camb. Philos. Soc.*
594   **73,** 157–180 (1998).

595   9.   Guillon, J., Julliard, R. & Leturque, H. Evolution of habitat-dependent sex
596   allocation in plants: superficially similar to, but intrinsically different from
597   animals. *J. Evol. Biol.* **19,** 500–512 (2006).

598 10. Dodson, C. H. Why are there so many orchid species□? *Lankesteriana* **7,** 99–103
599   (2003).

600 11. Fay, M. F. & Chase, M. W. Orchid biology: From Linnaeus via Darwin to the
601   21st century. *Ann. Bot.* **104,** 359–364 (2009).

602 12. Bateman, R. M. & Rudall, P. J. The good, the bad and the ugly: using naturally
603   occurring terata to distinguish the possible from the impossible in the orchid
604   floral evolution. *Aliso* **22,** 481–496 (2006).

605 13. Borba, E. L., Barbosa, A. R., Melo, M. C., Gontijo, S. L. & Oliveira, H. O.
606   Mating systems in the Pleurothallidinae (Orchidaceae): evolutionary and
607   systematic implications. *Lankesteriana* **11,** 207–221 (2011).

608 14. Pérez-Escobar, O. A., Gottschling, M., Whitten, W. M., Salazar, G. & Gerlach,
609   G. Sex and the Catasetinae (Darwin's favourite orchids). *Mol. Phylogenet. Evol.*
610   **97,** 1–10 (2016).

611 15. Pérez-Escobar, O. A. *et al.* Recent origin and rapid speciation of Neotropical
612   orchids in the world's richest plant biodiversity hotspot. *New Phytol.* **215,** 891–
613   905 (2017).

614 16. Dressler, R. L. *Phylogeny and classification of the orchid family*. **20,**
615   (Dioscorides Press, 1993).

616 17. Rudall, P. J. & Bateman, R. M. Roles of synorganisation, zygomorphy and
617   heterotopy in floral evolution: the gynostemium and labellum of orchids and
618   other lilioid monocots. *Biol. Rev* **77,** 403–441 (2002).

619 18. Rudall, P. J., Perl, C. D. & Bateman, R. M. Organ homologies in orchid flowers
620   re-interpreted using the Musk Orchid as a model. *PeerJ* **1,** e26 (2013).

621 19. Romero, G. A. in *Genera Orchidacearum: Vol . 5. Epidendroideae (part two)*
622   (eds. Pridgeon, A. M., Cribb, P. J., Chase, M. W. & Rasmussen, F. N.) 11–40
623   (Oxford University Press, 2009).

624 20. Chase, M. W. *et al.* An updated classification of Orchidaceae. *Bot. J. Linn. Soc.*
625   **177,** 151–174 (2015).

626 21. Ramirez, S. R. *et al.* Asynchronous diversification in a specialized plant-
627   pollinator mutualism. *Science* **333,** 1742–1746 (2011).

628 22. Pansarin, L. M., Castro, M. D. E. M. & Sazima, M. Osmophore and elaiophores
629   of *Grobya amherstiae* (Catasetinae, Orchidaceae) and their relation to
630   pollination. *Bot. J. Linn. Soc.* **159,** 408–415 (2009).

631 23. Darwin, C. *On the various contrivances by which british and foreign orchids are*
632   *fertilised by insects*. (Appleton and CO., 1877).

633 24. Pridgeon, A. M., Cribb, P. J., Chase, M. W. & Rasmussen, F. N. *Genera*
634   *Orchidacearum: Vol. 5. Epidendroideae (part two)*. (Oxford University Press,
635   2009).

636 25. Gregg, K. B. Sunlight-enhanced ethylene evolution by developing inflorescences
637   of *Catasetum* and *Cycnoches* and its relation to female flower production. *Bot.*
638   *Gaz.* **143,** 466–475 (1982).

639 26. Perez-escobar, O. A. Molecular phylogenetics, evolution of sexual systems and
640   historical biogeography of Darwin's favorite orchids (Catasetinae) and Swan

641   orchids (*Cycnoches* Lindl.). (Ludwig-Maximilians Universität, 2016).

642 27. Atsatt, P. R. & Rundel, P. W. Pollinator maintenance vs. fruit production:
643   partitioned reproductive effort in subdioecious *Fuchsia lycioides*. *Ann. Missouri*
644   *Bot. Gard.* **69,** 199–208 (1982).

645 28. Delph, L. F. Sex-differential resource allocation patterns in the subdioecious
646   shrub *Hebe subalpina*. *Ecology* **71,** 1342–1351 (1990).

647 29. Delph, L. F. & Wolf, D. E. Evolutionary consequences of gender plasticity in
648   genetically dimorphic breeding systems. *New Phytol.* **166,** 119–128 (2004).

649 30. Givnish, T. J. *et al.* Orchid phylogenomics and multiple drivers of their
650   extraordinary diversification. *Proc. R. Soc. B Biol. Sci.* **282,** 20151553 (2015).

651 31. Gravendeel, B., Smithson, A., Slik, F. J. W. & Schuiteman, A. Epiphytism and
652   pollinator specialization: drivers for orchid diversity? *Philos. Trans. R. Soc.*
653   *Lond. B. Biol. Sci.* **359,** 1523–1535 (2004).

654 32. Romero-González, G. A., Batista, J. & Bem, L. A Synopsis of the genus
655   *Cyrtopodium* (Catasetinae□: Orchidaceae). *Harvard Pap. Bot.* **13,** 189–206
656   (2008).

657 33. Whitten, W. M., Neubig, K. M. & Williams, N. H. Generic and subtribal
658   relationships in neotropical Cymbidieae (Orchidaceae) based on *matK/ycf1*
659   plastid data. *Lankesteriana* **13,** 375–392 (2014).

660 34. Pérez-Escobar, O. A., Balbuena, J. A. & Gottschling, M. Rumbling orchids: how
661   to assess divergent evolution between chloroplast endosymbionts and the nuclear
662   host. *Syst. Biol.* **65,** 51–65 (2016).

663 35. Tremblay, R. L., Ackerman, J. D., Zimmerman, J. K. & Calvo, R. N. Variation in
664   sexual reproduction in orchids and its evolutionary consequences: a spasmodic
665   journey to diversification. *Biol. J. Linn. Soc.* **84,** 1–54 (2005).

666 36. Hoorn, C. *et al.* Amazonia through time: andean uplift, climate change, landscape
667   evolution, and biodiversity. *Science* **330,** 927–931 (2010).

668 37. Antonelli, A., Nylander, J. A. A., Persson, C. & Sanmartín, I. Tracing the impact
669   of the Andean uplift on Neotropical plant evolution. *Proc. Natl. Acad. Sci. U. S.*
670   *A.* **106,** 9749–9754 (2009).

671 38. Antonelli, A. & Sanmartín, I. Why are there so many plant species in the
672   Neotropics? *Taxon* **60,** 403–414 (2011).

673 39. Hoorn, C. An environmental reconstruction of the palaeo-Amazon River system
674   (Middle–Late Miocene, NW Amazonia). *Palaeogeogr. Palaeoclimatol.*
675   *Palaeoecol.* **112,** 187–238 (1994).

676 40. Gentry, A. H. Neotropical floristic diversity: phytogeographical connections
677   between Central and South America, Pleistocene climatic fluctuations, or an
678   accident of the Andean orogeny? *Ann. Missouri Bot. Gard.* **69,** 557–593 (1982).

679 41. Pokorny, L. *et al.* Living on the edge: timing of Rand Flora disjunctions
680   congruent with ongoing aridification in Africa. *Front. Genet.* **6,** 1–15 (2015).

681 42. Alves, R. V. & Kolbek, J. Plant species endemism in savanna vegetation on table
682   mountains (*Campo Rupestre*) in Brazil. *Vegetatio* **113,** 125–139 (1994).

683 43. Antonelli, A., Verola, C. F., Parisod, C. N. & Gustafsson, a. L. S. Climate

684   cooling promoted the expansion and radiation of a threatened group of South
685   American orchids (Epidendroideae: Laeliinae). *Biol. J. Linn. Soc.* **100,** 597–607
686   (2010).

687 44. Freeman, D. C., Harper, K. T. & Charnov, E. L. Sex change in plants: old and
688   new observations and new hypotheses. *Oecologia* **47,** 222–232 (1980).

689 45. Zimmerman, J. K. Ecological correlates of labile sex expression in the orchid
690   *Catasetum viridiflavum. Ecology* **72,** 597–608 (1991).

691 46. Arditti, J. & Ghani, A. K. A. Numerical and physical properties of orchid seeds
692   and their biological implications. *New Phytol.* **145,** 367–421 (2000).

693 47. Yukawa, T., Ogura-Tsujita, Y., Shefferson, R. P. & Yokoyama, J. Mycorrhizal
694   diversity in *Apostasia* (Orchidaceae)indicates the origin and evolution of orchid
695   mycorrhiza. *Am. J. Bot.* **96,** 1997–2009 (2009).

696 48. Gentry, A. H. & Dodson, C. H. Diversity and biogeography of Neotropical
697   vascular epiphytes. *Ann. Missouri Bot. Gard.* **74,** 205–233 (1987).

698 49. Zotz, G. The systematic distribution of vascular epiphytes – a critical update. *Bot.
699   J. Linn. Soc.* **171,** 453–481 (2013).

700 50. Chomicki, G. *et al.* The velamen protects photosynthetic orchid roots against
701   UV-B damage, and a large dated phylogeny implies multiple gains and losses of
702   this function during the Cenozoic. *New Phytol.* **205,** 1330–1341 (2015).

703 51. Salazar, G. A., Cabrera, L. I., Gerlach, G., Hágsater, E. & Chase, M. W.
704   Phylogenetic relationships in *Mormodes* (Orchidaceae, Cymbidieae, Catasetinae)
705   inferred from nuclear and plastid DNA sequences and morphology. *Phytotaxa*
706   **263,** 18–30 (2016).

707 52. Monteiro, S. H. N., Selbach-Schnadelbach, A., de Oliveira, R. P. & van den
708   Berg, C. Molecular phylogenetics of *Galeandra* (Orchidaceae: Catasetinae) based
709   on plastid and nuclear DNA sequences. *Syst. Bot.* **35,** 476–486 (2010).

710 53. Pérez-Escobar, O. A., Parra-Sánchez, E. & Valdivieso, P. O. Inventario
711   orquideológico de la Reserva Bosque de Yotoco, Valle del Cauca. *Acta
712   Agronómica* **58,** 189–196 (2009).

713 54. Kolanowska, M., Pérez-Escobar, O. A., Parra-Sánchez, E. & Szlachetko, D. L.
714   *An illustrated field guide to the orchids of the Yotoco Forest Reserve (Colombia).*
715   (Wydawnictwo Uniwersytetu Gdańskiego, 2011).

716 55. Batista, J. N. *et al.* Molecular phylogenetics of Neotropical *Cyanaeorchis*
717   (Cymbidieae, Epidendroideae, Orchidaceae): geographical rather than
718   morphological similarities plus a new specie. *Phytotaxa* **156,** 251–272 (2014).

719 56. Maddison, W. P. & Fitzjohn, R. G. The unsolved challenge to phylogenetic
720   correlation tests for categorical characters. *Syst. Biol.* **64,** 127–136 (2015).

721 57. Davis, M. P., Midford, P. E. & Maddison, W. Exploring power and parameter
722   estimation of the BiSSE method for analyzing species diversification. *BMC Evol.
723   Biol.* **13,** 38–49 (2013).

724 58. Gregg, K. B. Variation in floral fragrances and morphology : incipient
725   speciation in *Cycnoches*? *Bot. Gaz.* **144,** 566–576 (1983).

726 59. Parra-Sánchez, E., Retana, J. & Armenteras, D. Edge Influence on diversity of

727   orchids in Andean cloud forests. *Forests* **7,** 1–13 (2016).

728 60. Carr, G. F. J. The genus *Cycnoches*: species and hybrids. *Orchid Rev.* 1–31
729   (2012).

730 61. Zachos, J. C., Dickens, G. R. & Zeebe, R. E. An early Cenozoic perspective on
731   greenhouse warming and carbon-cycle dynamics. *Nature* **451,** 279–283 (2008).

732 62. Neubig, K. M. *et al.* Generic recircumscriptions of Oncidiinae (Orchidaceae☐:
733   Cymbidieae) based on Maximum Likelihood analysis of combined DNA
734   datasets. *Bot. J. Linn. Soc.* 117–146 (2012).

735 63. Bone, R. E., Smith, J. A. C., Arrigo, N. & Buerki, S. A macro-ecological
736   perspective on crassulacean acid metabolism (CAM) photosynthesis evolution in
737   Afro-Madagascan drylands: Eulophiinae orchids as a case study. *New Phytol.*
738   **208,** 469–481 (2015).

739 64. Pérez-Escobar, O. A. *et al.* Andean mountain building did not preclude dispersal
740   of lowland epiphytic orchids in the Neotropics. *Sci. Rep.* **7,** 4919 (2017).

741 65. Irimia, R.-E., A. Pérez-Escobar, O. & Gottschling, M. Strong biogeographic
742   signal in the phylogenetic relationships of *Rochefortia* Sw. (Ehretiaceae,
743   Boraginales). *Plant Syst. Evol.* **301,** 1509–1516 (2014).

744 66. Bechteler, J. *et al.* Geographical structure, narrow species ranges, and Cenozoic
745   diversification in a pantropical clade of epiphyllous leafy liverworts. *Ecol. Evol.*
746   **7,** 638–653 (2017).

747 67. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software
748   version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30,** 772–
749   780 (2013).

750 68. Balbuena, J. A., Míguez-Lozano, R. & Blasco-Costa, I. PACo: a novel procrustes
751   application to cophylogenetic analysis. *PLoS One* **8,** e61048 (2013).

752 69. Feldberg, K. *et al.* A phylogeny of Cephaloziaceae (Jungermanniopsida) based
753   on nuclear and chloroplast DNA markers. *Org. Divers. Evol.* **16,** 727–742 (2016).

754 70. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more
755   models, new heuristics and parallel computing. *Nat. Methods* **9,** 772–772 (2012).

756 71. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-
757   analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313 (2014).

758 72. Ronquist, F. *et al.* Mrbayes 3.2: Efficient Bayesian phylogenetic inference and
759   model choice across a large model space. *Syst. Biol.* **61,** 539–542 (2012).

760 73. Miller, M. A. *et al.* A RESTful API for access to phylogenetic tools via the
761   CIPRES Science Gateway. *Evol. Bioinforma.* **11,** 43–48 (2015).

762 74. Erixon, P., Svennblad, B., Britton, T. & Oxelman, B. Reliability of Bayesian
763   posterior probabilities and bootstrap frequencies in phylogenetics. *Syst. Biol.* **52,**
764   665–673 (2003).

765 75. Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F. & Douzery, E. J. P.
766   Comparison of Bayesian and maximum likelihood bootstrap measures of
767   phylogenetic reliability. *Mol. Biol. Evol.* **20,** 248–254 (2003).

768 76. Drummond, A. J. & Bouckaert, R. R. *Bayesian evolutionary analysis with*
769   *BEAST 2*. (Cambridge University Press, 2014).

770 77. Bouckaert, R. *et al.* BEAST 2: a software platform for Bayesian evolutionary
771   analysis. *PLoS Comput. Biol.* **10,** 1–6 (2014).

772 78. Ramírez, S. R., Gravendeel, B., Singer, R. B., Marshall, C. R. & Pierce, N. E.
773   Dating the origin of the Orchidaceae from a fossil orchid with its pollinator.
774   *Nature* **448,** 1042–1045 (2007).

775 79. Poinar, G. Orchid pollinaria (Orchidaceae) attached to stingless bees
776   (Hymenoptera: Apidae) in Dominican amber. *Neues Jahrb. für Geol. und*
777   *Paläontologie* **297,** 287–293 (2016).

778 80. Töpel, M. *et al.* SpeciesGeoCoder: Fast categorisation of species occurrences for
779   analyses of biodiversity, biogeography, ecology and evolution. *BioRxiv* **9274,** 1–
780   15 (2014).

781 81. Givnish, T. J. *et al.* Phylogeny, adaptive radiation, and historical biogeography in
782   Bromeliaceae: Insights from an eight-locus plastid phylogeny. *Am. J. Bot.* **98,**
783   872–895 (2011).

784 82. Matzke, N. J. Model selection in historical biogeography reveals that founder-
785   event speciation is a crucial process in island clades. *Syst. Biol.* **63,** 951–970
786   (2014).

787 83. Chamberlain, S., Ram, K. & Hart, T. spocc: Interface to Species Occurrence Data
788   Sources. *R Packag. version 0.4.5. https//CRAN.R-project.org/package=spocc*
789   (2016).

790 84. Hijmans, R. J. & Elith, J. Species distribution modeling with R. (2016).

791 85. Arrigo, N. *et al.* A total evidence approach to understanding phylogenetic
792   relationships and ecological diversity in *Selaginella* subg. *Tetragonostachys*. *Am.*
793   *J. Bot.* **100,** 1672–1682 (2013).

794 86. Oksanen, J. *et al.* Package 'vegan'. *R Packag. ver. 2.0–8* (2013).
795   doi:10.4135/9781412971874.n145

796 87. Romero-González, G. A. & Nelson, C. E. Sexual dimorphism in *Catasetum*
797   orchids: forcible pollen emplacement and male flower competition. *Science* **232,**
798   1538–1540 (1986).

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

801 89. Wickham, H. *GGPLOT2: elegant graphics for data analysis.* (Springer-Verlag,
802   2009). doi:10.1007/978-0-387-98141-3

803 90. Gerlach, G. & Pérez-Escobar, O. A. Looking for missins swans: phylogenetics of
804   *Cycnoches. Orchids* **83,** 434–437 (2014).

805 91. FitzJohn, R. G. Analysing diversification with diversitree. *R Packag. ver. 0.9–2*
806   (2010).

807 92. Rabosky, D. L. & Goldberg, E. E. Model inadequacy and mistaken inferences of
808   trait-dependent speciation. *Syst. Biol.* **64,** 340–355 (2015).

809

810

811 **Acknowledgements**

824

825   Author contributions

826   O.A.P.E., G.C., F.L.C., J.V. and J.H. designed research; A.M and E.S. collected

827   samples; A.M., E.S. and O.P. performed the lab work; O.A.P.E., F.L.C., and G.C.

828   performed all analyses; O.A.P.E., G.C., F.L.C., J.V., B.K., J.H., G.G., A.M., and E.S.

829   wrote the manuscript.

830

831   Competing financial interest

832   The authors declare no competing financial interest.

1   **Tables**

2   **Table 1.** Best models and parameter values of BiSSE analysis on lineages with ESD

3   vs. alternative sexual systems, epiphytism vs. other plant habits, and male euglossine-

4   bee pollination vs. other pollination systems.

5

6   **Figures**

7   **Fig. 1.** Best scoring ML tree of Catasetinae (B) and sister subtribes (i.e. Eulophiinae,

8   Dipodiinae, Oncidiinae, Cyrtopodiinae, Cymbidiinae; A) obtained from non-

9   conflicting concatenated nuclear and plastid loci. Node pie-diagrams indicate

10   Bootstrap percentage values (MLBS > 75), where fully black diagrams indicate

11   MLBS of 100. Asterisks represent BPP values higher than 0.95. Representatives of

12   each clade in Catasetinae (except *Cyanaeorchis*) are shown in pictures [inset:

13   *Catasetum*: *Ca. ochraceum*; *Clowesia*: *Cl. russelliana*; *Cycnoches*: *Cyc. lehmannii*;

14   *Dressleria*: *D. eburnea*; *Galeandra*: *Ga. macroplectron*; *Grobya*: *G. galeata*;

15   *Mormodes*: *M. punctata*]. Pictures by: O. Pérez and G. Gerlach.

16   **Fig. 2.** Chronogram of Catasetinae and sister subtribes obtained under a relaxed clock

17   model, applied to a non-conflicting, concatenated nuclear and plastid loci. Minimum

18   and maximum age intervals are provided in Fig. S4. Time scale is provided in million

19   years (Mya). Node charts correspond to ancestral areas estimated under the

20   BAYAREA+J model. The LCA of Catasetinae is indicated with a black star and gains

21   of ESD are indicated with red circles. Coded biogeographical areas are colour-coded

22   following inset map, and are shown in front of taxa names (grey colours indicate taxa

23   distributed in more than one biogeographical region). The approximate time span of

24   Pebas and Acre mega-wetlands[36] is indicated by a grey rectangle. Inset: Coded areas

25   used for biogeographical analysis. Geopolitical boundaries map generated by

26   ArcMAP (http://www.sri.com) using political divisions and elevation data from

27   DIVA-GIS (http:www.diva-gis.org/data). Climate curve of the last 17 Myr

28   represented as a function of oxygen-isotope records[61].

29   **Fig. 3.** Evolution of ESD in Catasetinae and ancestral climatic-niche preferences

30   depicted as traitgrams obtained via ACE analyses of mean values of bioclimatic

31   variables. In all panels, coloured circles denote independent origins of ESD (colour-

32   coded by chronological order). Colour coded branches denote lineages estimated to

33   have ESD. Grey branches denote lineages inferred to have evolved alternative sexual

34   systems (i.e. no ESD). Lineages in which ESD has evolved are color-coded

35   accordingly. (A) Topological position of ESD origins (for posterior probabilities, see

36   Fig. S6) (B) Annual mean temperature (Bio 1); (C) Mean diurnal range (Bio 2); (D)

37   Isothermality (Bio 3); (E) Temperature seasonality (Bio 4); (F) Maximum

38   temperature of warmest month (Bio 5); (G) Maximum temperature of coldest month

39   (Bio 6); (H) Temperature annual range (Bio 7); (I) Mean temperature of wettest

40   quarter (Bio 8); (J) Mean temperature of driest quarter (Bio 9); (K) Mean temperature

41   of warmest quarter (Bio 10); (L) Mean temperature of coldest quarter (Bio 11); (M)

42   Annual precipitation (Bio 12); (N) Precipitation of wettest month (Bio 13); (O)

43   Precipitation of driest month (Bio 14); (P) Precipitation seasonality (Bio 15); (Q)

44   Precipitation of wettest quarter (Bio 16); (R) Precipitation of driest quarter (Bio 17);

45   (S) Precipitation of warmest quarter (Bio 18); (T) Precipitation of coldest quarter (Bio

46   19); (U) Altitude (meters). Note the little divergence through time of ancestral mean

47   bioclimatic variables between lineages having ESD, and those with alternative sexual

48   systems.

49   **Supplementary material**

50   **Appendix S1.** Extended Materials and Methods, supplementary tables and figures.

1 **Table 1**

Trait	Model	logL[1]	AICc[2]	λ[3] (0)	λ (1)	λ ratio (λ_1/λ_0)	μ[4] (0)
ESD	M_2	-248.854	506.035	1.253	2.024	1.614	-
Male euglossine-bee pollination	M_2	-240.045	488.416	1.827	2.138	1.174	-
Epiphytism	M_6	-253.720	517.931	1.276	1.56p	1.2507	-

2

3 [1]log Likelihood

4 [2]Akaike Informative Criterion corrected

5 [3]Speciation rates

6 [4]Extinction rates

7 [5]transition rates from state 0 to 1

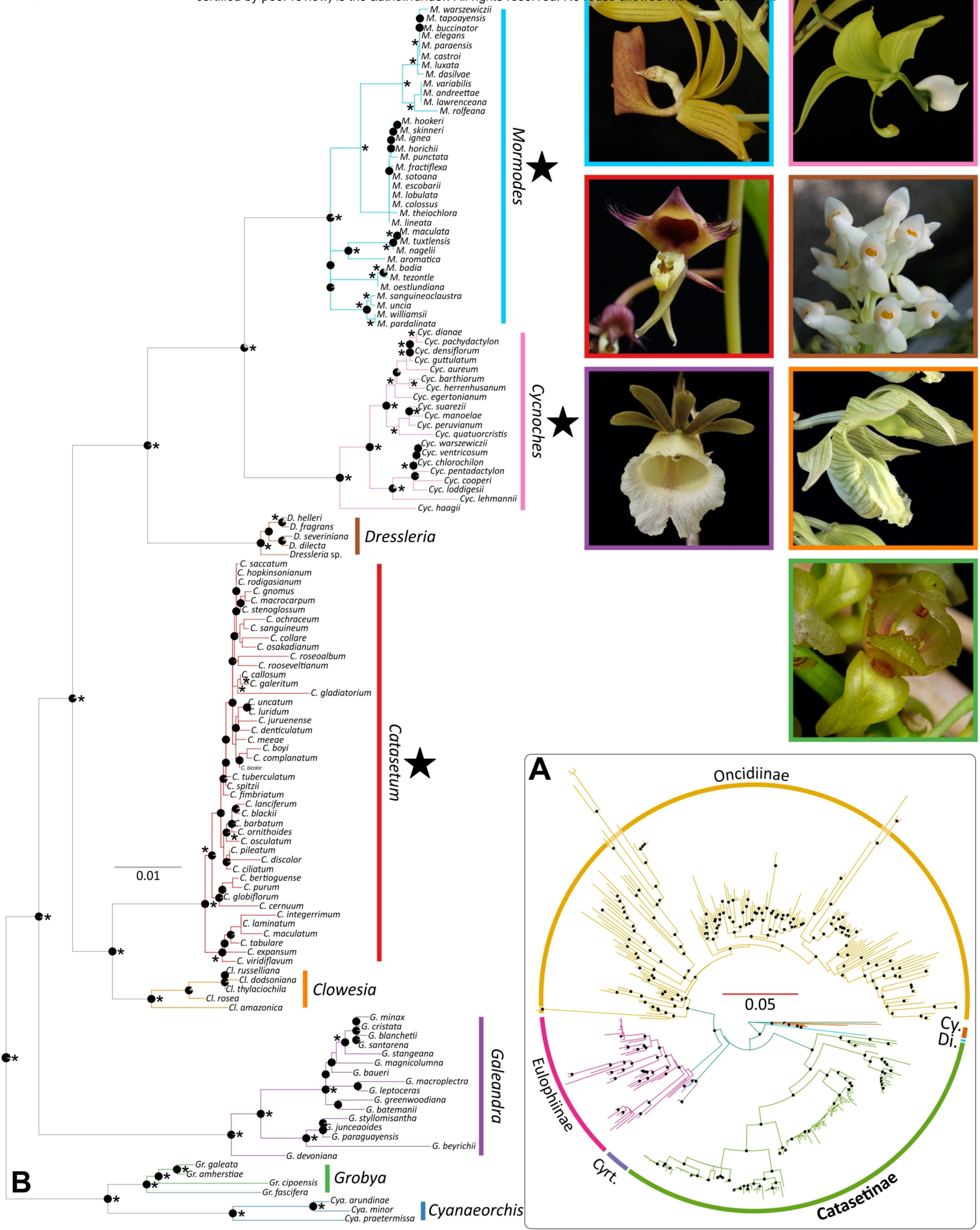8 [6]transition rates from state 1 to 0
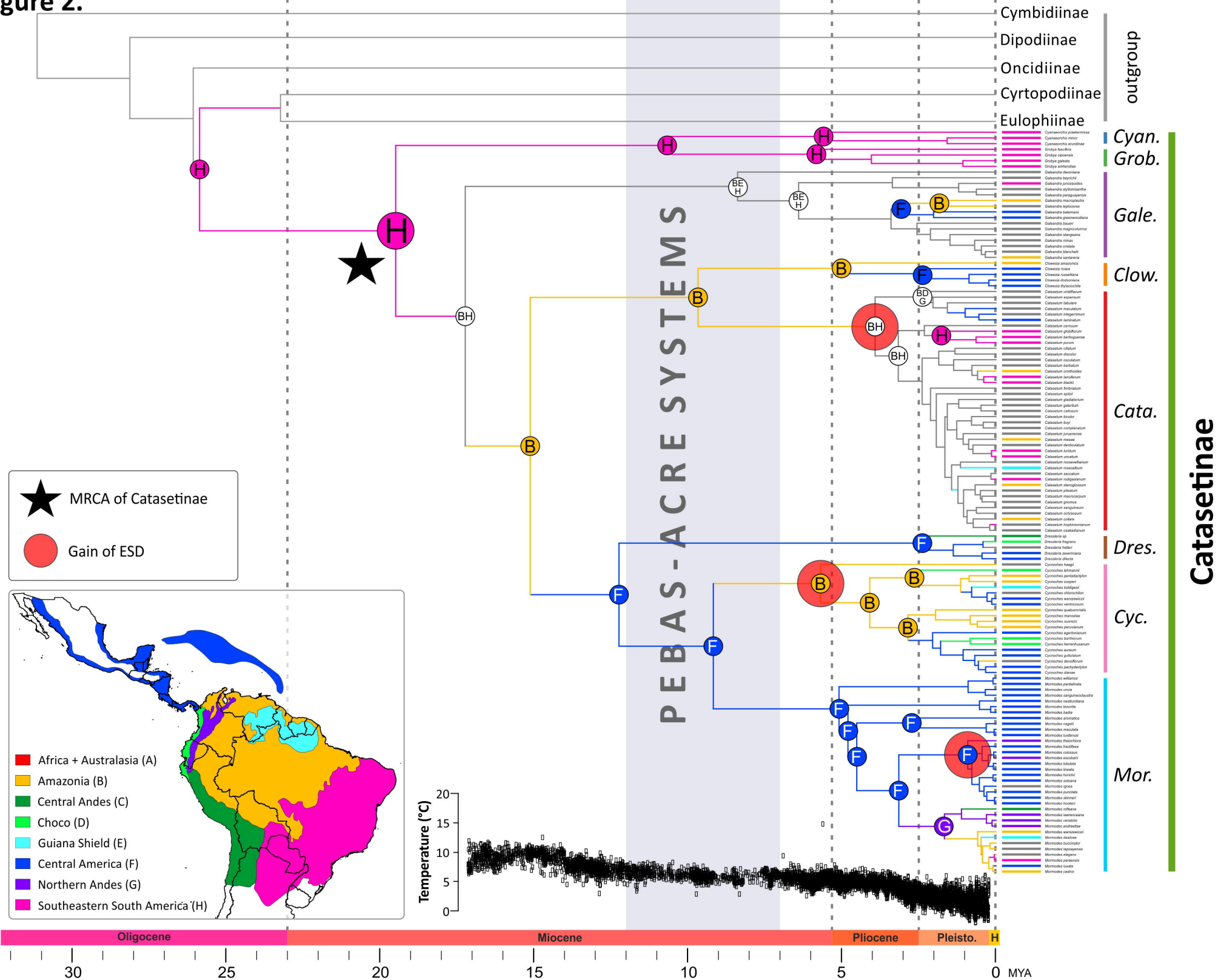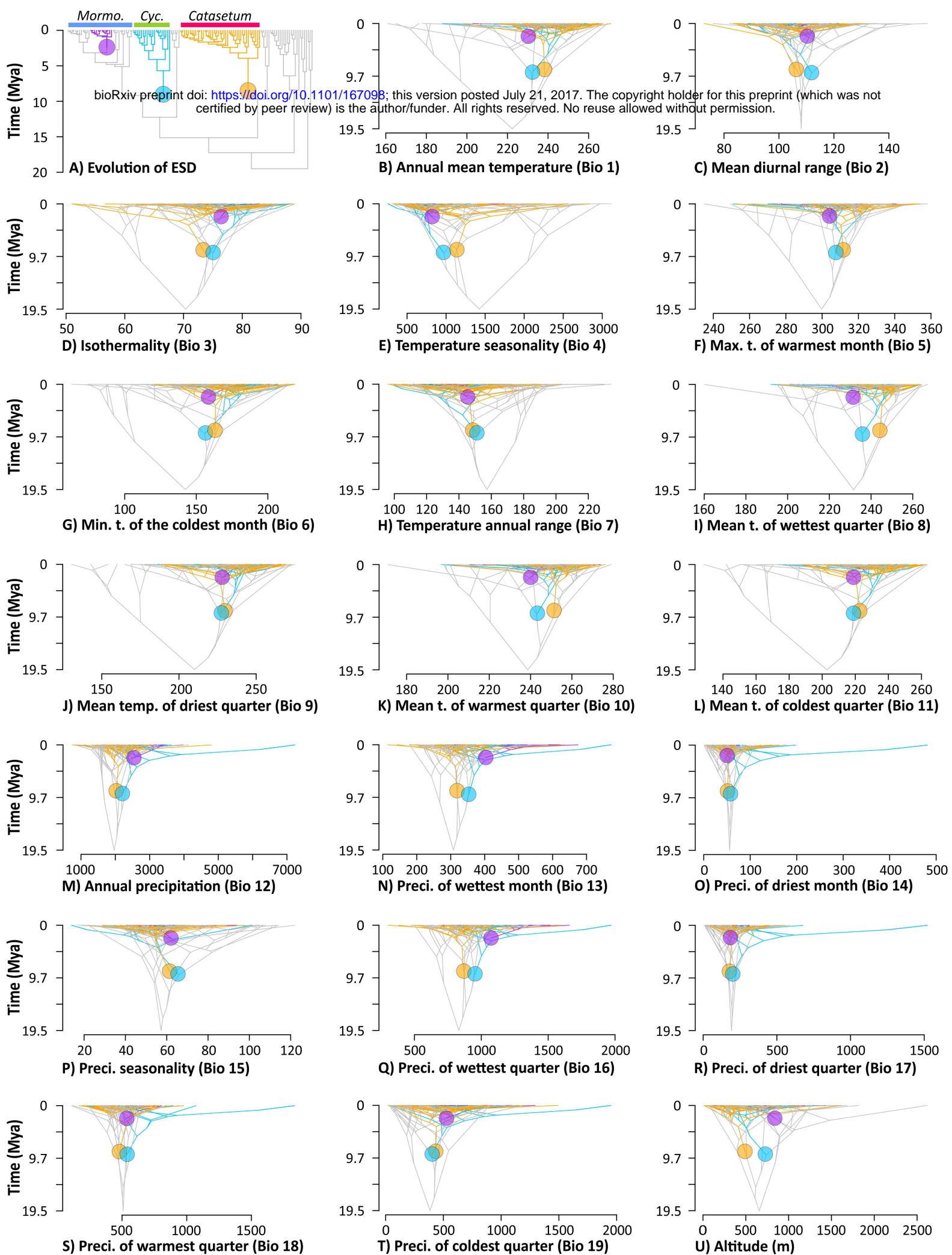
9

10
11

12

13

14

15

16

17

18

19

20

21

Figure 2.

A) Evolution of ESD

B) Annual mean temperature (Bio 1)

C) Mean diurnal range (Bio 2)

D) Isothermality (Bio 3)

E) Temperature seasonality (Bio 4)

F) Max. t. of warmest month (Bio 5)

G) Min. t. of the coldest month (Bio 6)

H) Temperature annual range (Bio 7)

I) Mean t. of wettest quarter (Bio 8)

J) Mean temp. of driest quarter (Bio 9)

K) Mean t. of warmest quarter (Bio 10)

L) Mean t. of coldest quarter (Bio 11)

M) Annual precipitation (Bio 12)

N) Preci. of wettest month (Bio 13)

O) Preci. of driest month (Bio 14)

P) Preci. seasonality (Bio 15)

Q) Preci. of wettest quarter (Bio 16)

R) Preci. of driest quarter (Bio 17)

S) Preci. of warmest quarter (Bio 18)

T) Preci. of coldest quarter (Bio 19)

U) Altitude (m)

Legends

● First gain of ESD

● Second gain of ESD

● Third gain of ESD

— *Catasetum*

— *Cycnoches*

— *Mormodes*