1 **Partial endoreplication stimulates diversification in the species-richest lineage of**
2 **orchids**

3 Zuzana Chumová^{1,2,6}, Eliška Záveská^{1,3,6}, Jan Ponert^{1,4,5,6}, Philipp-André Schmidt^{1,6}, Pavel
4 Trávníček^{*,1,6}

5
6 ¹Czech Academy of Sciences, Institute of Botany, Zámek 1, Průhonice CZ-25243, Czech Republic

7 ²Department of Botany, Faculty of Science, Charles University, Benátská 2, Prague CZ-12801, Czech Republic

8 ³Department of Botany, University of Innsbruck, Sternwartestraße 15, 6020 Innsbruck, Austria

9 ⁴Prague Botanical Garden, Trojská 800/196, Prague CZ-17100, Czech Republic

10 ⁵Department of Experimental Plant Biology, Faculty of Science, Charles University, Viničná 5, Prague CZ-
11 12844, Czech Republic

12

13 ⁶equal contributions

14 *corresponding author: Pavel.Travnicek@ibot.cas.cz

Abstract

Some of the most burning questions in biology in recent years concern differential diversification along the tree of life and its causes. Among others, it could be triggered by the evolution of novel phenotypes accelerating diversification in lineages that bear them. In the Pleurothallidinae, the most species-rich subtribe of plants on Earth with 46 genera and ~5,500 species, we constructed a completely new phylogeny and mapped on to it the type of endoreplication intending to trace how the phenomenon of partial endoreplication, which is unique to orchids, affects the differential diversification of lineages.

We have used NGS based target enrichment HybSeq approach for the reconstruction of the phylogeny and the flow cytometry to estimate the type of endoreplication. The BAMM and BiSSE analyses have been used to assess diversification rates and to trace the phenotype changes.

We have found that three of six changes in diversification rates are associated with changes in the endoreplication type and the clades bearing taxa with partial endoreplication showed higher net diversification rates. Our results demonstrate that multiple evolution of partial endoreplication within the subtribe considerably shapes the patterns of diversity and that partial endoreplication is a trait with an evolutionary significance.

Key words: orchids, diversification, partial endoreplication, evolution, Hyb-Seq, Pleurothallidinae

35 **Introduction**

36 Species richness is unequally distributed across the tree of life and remarkable differences are
 37 apparent even between closely related evolutionary lineages. In orchids, which are perhaps
 38 the most diverse plant family (rivalled only by the Asteraceae), there are several examples of
 39 disproportionally diversified species assemblages spanning from ancient monogeneric
 40 lineages (e.g. the tribes Codonorchideae, with two species, or the Thaisae, with a single
 41 species) persisting for dozens of millions of years undiversified [1,2] to clades with relatively
 42 recent massive diversification producing thousands of species (e.g. the largest plant subtribe
 43 Pleurothallidinae) [3]. The evolution of novel phenotypes triggering increased net
 44 diversification rates is often suggested as a driver of such disproportion, despite relatively
 45 rare evidence of such phenotypes. In orchids, accelerations of net diversification have been
 46 correlated with the evolution of pollinia, the epiphytic habit, CAM photosynthesis, tropical
 47 distribution and pollination by lepidopterans or euglossine bees [1]. Nevertheless, no such
 48 trait is confined to the orchids. By contrast, the phenomenon of partial endoreplication, so far
 49 known only from orchids, has never been tested for its diversification significance.

50 Endoreplication, the process of intraindividual polyploidization usually connected with cell
 51 differentiation, is unevenly distributed across the plant kingdom and is ubiquitously observed
 52 only in a few plant families such as the Brassicaceae or Cucurbitaceae (see Leitch &
 53 Dodsworth for an overview) [4]. Conventionally, endoreplication involves the doubling of
 54 the whole nuclear DNA content without cell division, resulting in the formation of
 55 endopolyploids. The exception from this rule are orchids, in which there exists so-called
 56 partial endoreplication (PE) [5,6], where a species-specific fraction of nuclear DNA is
 57 replicated, resulting in nuclei with (substantially) less than two-fold increase in DNA content
 58 [7]. It has to be noted that only one type of endoreplication has so far been observed in each
 59 species, so there is no evidence of any intra-specific switch from one type to another. Partial
 60 endoreplication is known from all subfamilies of orchids except one [8] and exhibits clade-
 61 specific behaviour with uneven distribution along the orchid phylogeny [7,9]. The biological
 62 significance of endopolyploidy remains elusive and, in plants, is usually attributed to the
 63 ability to grow in extreme or variable conditions [10,11]. Gene expression, enhanced by
 64 endoreplication and modulated under various environmental conditions, might be a factor that
 65 likely lies behind it [12]. Nevertheless, evidence that conventional and partial endoreplication
 66 differ in their biological significance is still lacking. All data gathered to date indicate a
 67 prevalence of PE in orchid subfamilies where it has once evolved [7], except for the species-

68 richest subfamily Epidendroideae, which includes ~80% of genera with exclusively the
69 conventional type of endoreplication (CE; our survey). Within the Epidendroideae, only the
70 monophyletic subtribe Pleurothallidinae combines exceptional diversification and almost
71 equal representation of PE and CE.

72

73 Pleurothallidinae is a young group restricted to tropical America which has recently
74 undergone an extraordinarily rapid diversification, resulting in more than 5,500 currently
75 accepted species [13,14]. Species of this group colonized various habitats from lowlands up
76 to nearly 4,000 m a.s.l. The Andean orogeny has been identified as an important factor in the
77 diversification of the Pleurothallidinae [3], and high variability in pollinator adaptation has
78 been hypothesized to also drive the speciation [15]. However, the specific factor generating
79 this diversification, which is exceptional among other Andean plant groups and is
80 asymmetric, remains elusive, and the question is whether propagation of partial
81 endoreplication within the subtribe might be such a factor.

82

83 To investigate this, we first constructed a robust phylogeny of the Pleurothallidinae based on
84 the Hyb-Seq approach utilizing orthologous low-copy nuclear (LCN) loci for phylogenetics
85 in non-model organisms [16]. Next, we dated the resulting species tree and reconstructed
86 ancestral states for both PE and CE to determine the shifts between them within the subtribe.
87 Simultaneously, we used the same tree to detect shifts in net diversification rates. Finally, we
88 compared the results to evaluate the association between diversification rates and the type of
89 endoreplication. Under the assumption of such an association, we tested the hypothesis that a
90 shift in the type of endoreplication precedes each shift in diversification rates and that clades
91 with partial endoreplication exhibit greater diversification rates.

92

93 **Material & Methods**

94 *Taxon sampling*

95 We sought to sample as many morphological groups and previously recognized lineages as
96 possible. Because living tissues are required for the estimation of the type of endoreplication,
97 we were unable to use herbarium specimens and were limited by the availability of living
98 plants in our collection. Finally, we sampled 40 out of 46 recognized genera that were
99 represented by 341 specimens of 337 species, plus seven species of outgroups. Because the
100 generic concept of Pleurothallidinae changes relatively frequently, we mostly follow the

concept of Karremans [13]. *Kraenzlinella* and *Dondodia* are included in *Acianthera* following Karremans et al. [17]. *Trichosalpinx* is treated in its strict sense following Bogarín et al. [18]. Morphologically unusual species, which have sometimes been treated as monospecific genera but without molecular phylogenetic evidence (*Andreettaea*, *Madisonia*), are classified as synonymous with *Pleurothallis*, following WCSP [14]. The adopted concept of genera of the Pleurothallidinae supplemented with the number of recognized species [14] is provided in Table S1, the complete list of analysed species in Table S2.

DNA extraction

Total genomic DNA was extracted from 0.5 g of silica-dried leaf material via the Sorbitol method [19] with two modifications: 1,600 µl of Extraction Buffer per sample were used (instead of 1,300 µl) and 6 µl of RNase instead of 4 µl in the first step. The quality of the DNA was checked on 1% agarose gel and using a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA).

Probe design

To select loci that are nuclear, single-copy and have orthologs across the Pleurothallidinae for sequence capture (HybSeq), we utilize our genomic data for *Stelis pauciflora* (P022) and the transcriptome for *Masdevallia yuangensis* (ID JSAG) from the 1,000 Plants (1KP) initiative (<http://www.onekp.com/samples/single.php?id=JSAG>). For these two species from different genera of Pleurothallidinae, we expected to find loci that will have likely orthologs also throughout the entire subtribe. To select loci, we utilized the marker development pipeline Sondovac [16] as described in detail in Results S1. One probe set targeting a total of 4,956 nuclear DNA loci was prepared using MYbaits (MYcroarray Ann Arbor, MI, USA). Biotinylated RNA probes were 120 bp long with 2× tiling density over target sequences. Additional checks were performed to eliminate probes targeting multi-copy loci by probe manufacturer.

HybSeq Library preparation

For each sample, an enriched genomic library was prepared following the protocol of manufacturer as described in detail in Results S1. Twenty-four samples per library were pooled in equimolar ratios (overall, 18 libraries were prepared). In-solution sequence capture of 4,956 exons from 1,200 nuclear low-copy genes designed as described above was done using MYbaits custom probes (MYcroarray, Ann Arbor, MI, USA) following the

135 manufacturer's protocol (see details in Results S1). Hyb-Seq libraries were sequenced on an
136 Illumina MiSeq instrument at OMICS Genomika (DNA sequencing laboratory of the
137 Biological Section of Charles University), Biocev, Vestec, using 300-cycle kit (v.3, Illumina,
138 Inc., San Diego, CA, USA) to obtain 150 bp paired-end reads.

139

140 *Processing of raw reads and species tree reconstruction*

141 We used the HybPhyloMaker v.1.6.4 pipeline [20] for raw read filtering, mapping of the
142 filtered reads to a pseudoreference (exon sequences used for probe design, separated by
143 strings of 400Ns) and construction of gene alignments (see details in Results S1). Also gene
144 trees and the species tree were reconstructed according to the HybPhyloMaker pipeline. Gene
145 alignments with more than 30% of missing data and less than 100% species presence were
146 excluded from further analyses. Alignment characteristics were calculated using AMAS [21],
147 trimAl 1.4 [22] and MstatX (github.com/gcollet/MstatX). Gene trees were reconstructed
148 using RAxML 8.2.4 [23] with 1,000 rapid bootstrap replicates (BS) and the GTRGAMMA
149 model. Summary statistics were generated using AMAS [21] and plotted using R 3.2.3 [24].
150 The multispecies coalescent model implemented in ASTRAL 5.6.1 [25] was used to construct
151 species trees with default settings. For usage in subsequent analysis (PhyParts), an additional
152 species tree was constructed with ASTRAL 5.6.1 from gene trees with < 50% BS collapsed
153 branches into a polytomy created in TreeCollapserCL 4 [26]. Additionally, alignments of all
154 low-copy genes were concatenated and analysed by taking the supermatrix approach using
155 FastTree 2 [27]. An overview of the reads obtained for all taxa involved is given in Table S2.

156

157 *Gene tree species tree (in)congruence*

158 We evaluated the level of concordance among gene trees (with a maximum of 30 percent of
159 missing data and with collapsed nodes with < 50 % support) against the ASTRAL species
160 tree with the program PhyParts (<https://bitbucket.org/blackrim/phyParts>) [28]. The species
161 tree was previously rooted in Figtree [29] with representatives from subtribe Laeliinae as an
162 outgroup. The output obtained with PhyParts was visualized by plotting pie charts on the
163 ASTRAL species tree and ML concatenated tree with the script PhyPartsPieCharts
164 (<https://github.com/mossmatters/MJPythonNotebooks>) using the ETE3 Python toolkit [30].

165

166 *Temporal scaling of the species tree*

167 A few unambiguous orchid macrofossils are available for Orchidaceae, e.g. [31,32], but these
168 are assigned to lineages very distantly related to the Pleurothallidinae. Because using distant

outgroups to calibrate our phylogenies would have created extensive sampling heterogeneities, which can result in spurious results [33], we had to rely on secondary calibrations. As secondary calibration points we relied mainly on a previously fossil-calibrated Orchidaceae-wide phylogeny; in particular we used diversification time of the subtribe Pleurothallidinae to be ca 20 million years ago [3]. For the large genomic datasets similar to ours, the use of Bayesian methods for dating analyses (e.g. BEAST) [34] is frequently unfeasible, as they require a long computational times [35–38]. Here we used the RelTime method [39], a fast-dating and high-performance algorithm that is implemented in MEGA X [40]. We used a species tree rooted with representatives of the subtribe Laellinae as a starting tree and concatenated alignment of full sequences of all 168 nuclear loci as source data for branch length estimation. The GUI-based version of RelTime in MEGA X was utilized, using the RelTime-ML option for timetree estimation and GTR with four discrete gamma rate categories as a model for the sequence alignment. As a time constraint, we used prior information on the diversification of the Pleurothallidinae, specifically 20 ± 5 Mya [3].

Endoreplication mode estimation

Data on endoreplication mode (partial vs conventional endoreplication) were estimated via flow cytometric analyses of all samples included in the tree according to the methodology described in detail by Trávníček et al. [8]. This approach is based on the quantification of the amount of DNA that has undergone endoreplication in the course of cell differentiation. Whereas plants with 100% replicated DNA (endoreplicated nuclei of subsequent sizes always differ two-fold) are assigned to the conventional type of endoreplication, plants with less than 100% replicated DNA (endoreplicated nuclei of subsequent sizes always differ by less than two-fold) are assigned to partial endoreplication (Fig. S1). The arbitrary threshold for assignment was set to 95% of replicated DNA because of possible measurement errors. The assignment of all included taxa was conducted via FloMax v2.4d (Partec GmbH) and is given in Table S2. In addition to the assignment of all involved ingroup plants (341 individuals), 340 additional plants randomly distributed along an ingroup clade were assigned to one of the endoreplication types to increase the precision of the estimation of the proportion between the two types of endoreplication within the Pleurothallidinae.

199 *Ancestral state reconstruction*

200 Using the species tree, we reconstructed ancestral states for endoreplication type under
 201 a BiSSE model (binary state speciation and extinction model) [41] implemented in the R
 202 package *diversitree* [42]. Incomplete sampling was accounted for by a state-dependent
 203 sampling factor that was estimated by the use of an extended dataset (see above). We started
 204 with a full six-parameter model and performed model simplification in a maximum likelihood
 205 framework based on likelihood ratio tests (implemented in the R package *lmtest*) [43]. Two
 206 best-fitting five-parameter BiSSE models with equal likelihood were found: the first model
 207 with single extinction rate and the second with a single speciation rate. Initially, the
 208 reconstruction using stochastic character mapping was performed using the R package
 209 *phytools* v.06-99 [44] and the function `make.simmap`. This method fits a continuous-time
 210 reversible Markov model for the evolution of the trait and then simulates stochastic character
 211 histories using that model and the tip states on the tree (sensu Bollback) [45]. We simulated
 212 1,000 character histories for the endoreplication trait and summarized it (appropriate nodes in
 213 Fig. 1; all nodes in Fig. S2). We also performed ancestral-state reconstruction for the BiSSE
 214 models (Figs. S3–5).

216 *Diversification rates*

217 We used Bayesian analysis of macroevolutionary mixtures, BAMM v2.5.0 [46] to estimate
 218 the pattern of diversification within the Pleurothallidinae. The model allows for variation in
 219 rates over time and among lineages, and despite criticism that it uses an incorrect likelihood
 220 function and prior settings [47], it provides a robust and appropriate tool for estimating
 221 lineage diversification [48,49]. BAMM allows for the implementation of incomplete clade-
 222 specific taxonomic sampling, which was necessary because of the use of a phylogeny
 223 reflecting just a fraction of the species diversity of the Pleurothallidinae (341 individuals in
 224 the phylogeny vs ~ 5,500 recognized species). We therefore applied fractioning of recognized
 225 diversity according to clade membership and virtually attribute each tip within a clade to
 226 represent the proportional fraction of known clade species richness (Table S1). We performed
 227 multiple BAMM runs of 50 millions generations, sampling every 2,000 generations, and
 228 checked the convergence using the R package *coda* v0.19-2 [50]. The effective sample size
 229 (ESS) of the log-likelihood and the number of shift events present in each sample remarkably
 230 exceeded the recommended threshold of 200. The analysis was repeated three times to ensure
 231 the convergence of all runs. BAMM output data were further processed using the R package
 232 *BAMMtools* v2.1.7 [51] that enables complex analysis and graphical representation.

233

234 *Diversification vs trait correlates.*

235 We estimated speciation, extinction and net diversification rates for clades according to the
 236 type of endoreplication – PE, CE and mixed clades. We extracted posterior distribution of
 237 state-dependent rates from BAMM analyses and examined the differences in rates between
 238 clades. For this we used the getCladeRates function of *BAMMtools* to obtain mean rates for
 239 particular clades. We compared state-dependent rates via credible intervals of differences to
 240 assess their significance. Similarly, we examined the best BiSSE model (with equal
 241 extinction rates) using a MCMC chain of 10,000 generations. We set the exponential prior to
 242 $1/2r$, where r was the character-independent diversification rate inferred from the BiSSE
 243 starting point. We removed 10% a burn-in and summarized MCMC samples to assess the
 244 variation in state-dependent rates. Finally, we calculated differences in posterior distributions
 245 of rates for character-states (PE vs CE) and estimated their credible intervals for the best (i.e.
 246 that having equal extinction rates) and the full BiSSE model (Fig. S6). In the last step, we
 247 compared the results derived from the BAMM and BiSSE approach.

248

249 *Relative flower size*

250 We recorded simple morphological traits (length of the ramicaul, length and width of leaves,
 251 and length of sepals and petals) from the appropriate literature sources to estimate the relative
 252 size of generative organs (flowers) to that of vegetative part of plants. We excluded taxa with
 253 unclear species determination. Those traits were investigated with respect to the type of
 254 endoreplication to elucidate the possible impact of PE on the morphological configuration of
 255 plants under investigation. We gathered data for 322 taxa (Table S2) and analysed them by
 256 ANOVA with and without phylogenetic independent contrasts (PIC) in the R package
 257 *phytools* [44] and *stats*, respectively.

258

259 **Results**

260 *Phylogenetic and dating analyses*

261 Species tree topologies obtained via the ASTRAL method and the supermatrix approach
 262 (concatenation) based on 168 loci differed only slightly, both showing high support values of
 263 Bayesian posterior probabilities (PP) in the case of the ASTRAL tree and SH-aLRT (SH) in
 264 the case of the supermatrix approach (Fig. S7). The large size and complexity of genomic
 265 datasets for phylogenetic reconstruction introduce significant phylogenetic noise and conflict

into subsequent analyses. For such datasets the Bayesian posterior distributions, SH-like local supports and even bootstrap supports are less informative because usually only a small minority of nodes show lower than maximum support, and the values might be biased by the size of the supermatrix [28,52]. Therefore we also examined the number of gene trees that are in concordance and in conflict with the species tree topology to interpret the support for particular phylogenetic units as well as to identify potential sources of conflicts like hybridization or incomplete lineage sorting [28]. PhyParts analysis of 168 gene trees and the ASTRAL species tree (Fig. S7a) shows high support (more than 50% of gene trees, hereafter GT, that support the species tree topology) for monophyletic Pleurothallidinae (160 GT) as well as for the core of the subtribe after separation of the basal genera *Dilomilis* and *Neocogniauxia* (144 GT). Support for particular genera and groups of genera, however, differs markedly across the phylogeny as described thoroughly in Methods S1. In addition, the groups of our interest, that is, the genera where shifts in diversification rates were observed (see below), differed in the number of gene trees that supported their monophyly, although the posterior probability was in all cases 1. In particular, the genus *Lepanthes* had the highest proportion of concordant/conflicting gene trees – 147/1. In the genera *Pleurothallis*, *Dracula* (excluding the potentially hybridogenous species *D. xenos*, see details in Methods S1) and *Stelis*, proportions of concordant genes were lower but still greater than those of conflicting ones (90/12, 60/29 and 40/32, respectively) and for the genus *Masdevallia* (including one species of the polyphyletic genus *Diodonopsis*) the proportion of conflicting gene trees exceeds that of concordant ones (28/47). The first clade that includes the genus *Masdevallia* and has a greater proportion of concordant genes (65/29) also covers the genera *Porroglossum* and *Diodonopsis*.

The absolute ages obtained for Pleurothallidinae chronograms are in agreement with previously published dated phylogenies, e.g. [3,53–55]. Divergence time estimates and 95% credibility intervals (CIs) inferred for all nodes of Pleurothallidinae chronograms are presented in Fig. S8. For the sake of clarity, we present the results using a simplified dated species tree visualizing major clades that roughly correspond to the accepted concept of genera. We divided clades with the incidence of paraphyletic genera into monophyletic lineages and supplemented the genus names with numbers to distinguish them (Fig. 1). Whereas the divergence time of the ingroup (Pleurothallidinae) was set as prior to 20 mya (± 5 mya), the age of the subtribe after the separation of the basal genera *Dilomilis* and *Neocogniauxia* (i.e. diversification of the Pleurothallidinae) was estimated to be 15.8 (CI 9,

24) mya. Below we present divergence times of groups of our interest along with information about their diversification rates.

Modelling diversification rates

To obtain diversification patterns within the phylogenetic tree, we used Bayesian analysis of macroevolutionary mixtures (BAMM) [46]. The analysis revealed multiple shifts in diversification rates in the Pleurothallidinae (Fig. 1, Fig. S9). We found strong support for 4–11 remarkable shifts in diversification rates across the subtribe, with a cumulative posterior probability of 0.98. The best models generated by Bayesian credible sets of shifts consistently show the six most preferred shifts (Fig. 1, Fig. S10). The first shift is located at the basis of the subtribe after the separation of the basal genera *Dilomilis* and *Neocogniauxia* (ca 15.8 mya), the other five shifts are more or less confined to the genera *Lepanthes*, *Dracula*, *Masdevallia*, *Pleurothallis* and *Stelis*, respectively. Whereas the model identified *Lepanthes* and *Masdevallia* as participating in increased diversification as whole genera, other genera are involved partially. The genus *Dracula* without the controversially placed *D. xenos*, the genus *Pleurothallis* without numerous basal lineages and genus *Stelis* only in its nominate subgenus *Stelis* (Fig. S9). Mean divergence times of the clades with increased diversification rates are generally not older than 4 million years (see details in Table S3), where group of *Masdevallia* 1 & 2 (incl. *Diodonopsis* 2) seems to be the oldest one (mean 3.8 mya, 95% CI 1–7.6 mya). In accordance with common practice [56], we also computed the branch-specific marginal probability and marginal odds (branch-specific pseudo-Bayes factors) to visualize the branches with rate shift locations. Both approaches show consistent results that are in accordance with the location of preferred shifts (Fig. S11). We also performed an analysis of diversification rates over time to reveal clade-specific rates in contrast to rates of the background (the rest of the Pleurothallidinae excluding the particular clade; Fig. 2). Those data show contrasting patterns in net diversification rates, diverging by the type of endoreplication. Whereas rates for clades with partial endoreplication (Fig. 2c and 2d) increase constantly, in accordance with the background, rates for clades with conventional endoreplication (Fig. 2e and 2f) exhibit a decrease. The exceptionally high net diversification rate for *Lepanthes* (Fig. 2b), the only diversified clade with mixed endoreplication, is constant over time.

Ancestral state reconstruction and diversification

We used ancestral state reconstruction via BiSSE and stochastic mapping to trace the possible shifts in the type of endoreplication and their possible coincidences with shifts in diversification rates. Assuming the significance of endoreplication in the diversification of orchids, we hypothesized that a shift in the type of endoreplication type preceded a shift in diversification rates. Such a pattern is traceable namely in the clade *Draconanthes/Lepanthes*, where massive diversification (Fig. 2b) immediately follows the shift to partial endoreplication or to a mixture of both types (Fig. S2). Other shifts of diversification rates in the *Dracula* and *Masdevallia* clades follow the shift to exclusively partial endoreplication with a noticeable delay (Fig. S2). The remaining increases of net diversification in *Pleurothallis* and *Stelis* do not show any association with shifts in endoreplication type because they are confined to clades with exclusively CE (Fig. 1, Fig. S2).

Trait-dependent diversification rates

Finally, we estimated trait-dependent diversification rates in clades with differential presence of endoreplication types. We primarily used the approach of taking posterior distribution of state-dependent rates from BAMM analyses. The clades were divided into three groups according to endoreplication type: CE clades containing only taxa with the conventional type, PE clades with exclusively the partial type and mixed clades where taxa exhibit either the conventional or the partial type of endoreplication. The clades were treated at the level of genera, but reflecting monophyletic lineages with the exception of the genus *Acianthera*, which was a priori divided down to the basalmost lineage corresponding to the subgenus *Kraenzlinella* (PE clade) and the rest of species (CE clade). Under this treatment we revealed significant shifts towards increased net diversification and speciation rates in PE and mixed clades; however, extinction rates remained almost the same for all clades (Fig. 3). Significance was assessed by examining the posterior distribution of differences between the clades (Fig. S12). Alternatively, we used posterior probabilities from the fit of the BiSSE model, where only two states of the trait were recognized (PE vs CE) because the analysis was carried out at the species level. The BiSSE model with equal extinction rates shows an even more prominent shift of net diversification rates for taxa with PE (Fig. S13). Additionally, we tried two alternative setups of clade partitioning for the interpretation of state-dependent rates from BAMM. We analysed the data at the level of whole clades (without dividing the genus *Acianthera*) and, conversely, at the finest (species) scale, where the assignment to mixed clades was limited only to sister taxa with a different type of

endoreplication. Whereas the first approach changed noticeably, the estimation of net diversification rates for mixed clades, the second approach provided similar results as previous analyses based on BAMM and BiSSE data (Fig. S14).

Morphological traits

We also tested the possible impact of PE on the phenotype to trace the importance of saving energy in connection with morphology. For this, we made a simple comparison of vegetative and generative parts of the Pleurothallidinae within our phylogeny and analysed data with respect to the type of endoreplication (Table 1). Flower size (represented by the length of the sepals), in both an absolute and relative way, was singled out as the most prominent difference between taxa differing in the type of endoreplication even when the data were tested under the phylogenetic frame.

Discussion

We present here the first thorough phylogenetic study of the whole Pleurothallidinae based on NGS data, using own designed probes for target enrichment according to [16]. Our approach revealed relatively high support for the current taxonomic concept in many lineages (genera), but indicated also paraphyly in some salient examples such as *Masdevallia*, *Diodonopsis*, *Tubella*, *Phloeophila*, *Specklinia*, *Platystele* and *Echinosepala*. The latter findings are either new or contradict the latest phylogeny of the subtribe inferred by [3] based on ITS and *matK*. We also present a relatively critical approach to the interpretation of phylogenies based on genomic data while taking into account numbers of gene trees that are concordant and discordant with the species tree reconstruction. We see multiple indications of hybridogenous origins of species or entire lineages (e.g. species of the genera *Myoxanthus*, *Restrepia*, *Zootrophion*, *Tubella*, *Lepanthopsis*, *Dracula*, *Specklinia*, *Stelis* and several others). Reticulate evolution has been documented also in several other orchid groups like *Paphiopedilum* (subfamily Cypripedioideae) [57], *Dactylorhiza* (subfamily Orchidoideae) [58] and *Polystachya* (subfamily Epidendroideae) [59], which may indicate that this process plays an important role in orchid evolution. In parallel we see indications of rapid radiations accompanied by massive gene tree – species tree incongruence likely caused by incomplete lineage sorting (e.g. in the genera *Masdevallia* and *Stelis*). Similarly high proportions of conflicting gene trees have been observed in young lineages of the rapidly diversified genus *Picris* [60]. We believe that our approach to thoroughly evaluating patterns of gene trees that

398 conflict or support the topology of a species tree will be valuable for further detailed
399 systematic/taxonomic studies, where the above mentioned evolutionary phenomena need to
400 be taken into account, as shown, for example, by [61].

401 Possible weaknesses of our research stem from the nature of the subtribe under study and the
402 impossibility to account for all recognized species, of which there are more than 5,500
403 [13,14]. Despite low coverage of the subtribe in our phylogeny, accounting for ~6.2%, we are
404 convinced of the reliability of our results because we covered almost all out of the 46
405 recognized genera (except for six very small genera; Table S1). A very similar approach
406 including a near-complete generic level encompassing only a fraction of species was taken by
407 Helmstetter et al. [62]. Our conviction is supported by convergent results from two
408 completely independent methods with different rationales and handling with incomplete
409 sampling. In addition, our phylogeny fulfils basic recommendations for BiSSE-like analyses
410 to avoid the main pitfalls of low sample size and strong tip ratio bias [63,64].

411
412 The time coincidence of massive acceleration in the five most diverse lineages (~2.4–3.8 Ma)
413 points to the existence of a common, probably an environmental, trigger. The establishment
414 of suitable habitats after the main orogenesis of the Andes has been postulated as the most
415 reliable explanation [3]. However, clade-specific net diversification rates exhibit substantial
416 differences (Fig. 2), suggesting the existence of other factors differentiating the clades. Under
417 this assumption, we tested the possible impact of partial endoreplication on divergence within
418 clades. We have found three of five shifts in diversification rates related to genera associated
419 with a shift from conventional to partial endoreplication. In addition, our data unequivocally
420 imply an association between greater diversification rates and PE clades (Figs. 2 and 3). So,
421 how might partial endoreplication promote diversification in orchids? PE might be enhanced
422 by the necessity to have an increased gene expression, crucial for plants with fast expansion
423 growth of cells. Such growth is not attained by increasing the ploidy level of nuclei, but by
424 regulating the expression of cell wall-modifying genes that allow cell growth via vacuolar
425 expansion [65]. Therefore, endoreplication might be of particular importance for tissues
426 containing extremely rapidly expanding cells, where an increase in the gene copy number
427 through endoreplication might be a way to cope with the high demand for new cell wall
428 material. The possible advantage of PE over CE might lie in its relatively low energetic and
429 supply costs because increased expression can be facilitated by endoreplication of just
430 a fraction of the genome [9]. This might be important especially in extreme environmental

conditions or under strong competitive pressure, that is, in places the majority of the Pleurothallidinae grow. We can further speculate that the energy saved by partial endoreplication might be invested into the formation of more attractive flowers for pollinators, so PE might serve as a trigger of increased diversification via speciation driven by pollinator adaptation [15]. According to this assumption, genera with the largest flowers within the subtribe (*Dracula* and *Masdevallia*) perform only PE whereas genera known for relatively large plants with small flowers (*Myoxanthus*, *Octomeria*, *Pleurothallis*, *Stelis*) perform only CE. Moreover, several other genera with a high incidence of PE are tiny plants, where the size of the flower is relatively large compared to that of vegetative parts of the plant (*Anathallis*, *Brachionidium*, *Lepanthes*). We therefore decided to analyse relative proportions of flowers to vegetative parts in sampled species of the Pleurothallidinae, and the results of this analysis clearly show that species with PE can invest more resources into producing larger flowers (Table 1). Even though the role of large flower size in myophilous (i.e. fly-pollinated) species, a substantial fraction of the Pleurothallidinae, remains a mystery, it has been observed in the genera *Restrepia* and *Telipogon* that osmophores can be distributed on various parts of flowers, sometimes very far from each other, likely to produce a more intensive gradient of a scent or to make ‘scent trails’ to guide pollinators [66,67]. We could therefore hypothesize that large flowers might provide an evolutionary advantage also in myophilous species of the Pleurothallidinae. However, there is another completely different, yet also plausible, explanation for the dominant diversification of PE clades. Endoreplication is induced in mycorrhizal tissues by fungal accommodation in arbuscular mycorrhizas [68] as well as in orchid mycorrhizas [69,70]. Taking into account that orchids are completely dependent on nutrition mediated by mycorrhizal fungi at the (post-) germination stage [71,72], we can deduce that they are predetermined to deal with endoreplication. We can therefore speculate that partial endoreplication might be a way to tackle the predetermination fostering its advantages while minimizing its disadvantages by saving resources also during mycorrhiza-induced endoreplication. Taking into account that PE is prominent in orchid species with larger genome sizes [7], we can further speculate that PE might represent an effective evolutionary response on how to simultaneously cope with limited sources and endopolyploidy. Seen in this light, the greater diversification rate in clades harbouring PE makes perfect sense. Anyway, elevated diversification rates are not exclusively associated with shifts in endoreplication type within the Pleurothallidinae, but clades with partial endoreplication exhibit greater rates in direct comparison to clades with conventional endoreplication (Fig. 3). To sum up, we can conclude that PE plays a noticeable

role in the evolution of the Pleurothallidinae and contributes to the formation of the species-
richest subtribe of plants on Earth.

Conclusions

Our results demonstrate that the largest orchid subtribe Pleurothallidinae diversified extremely rapidly in the last ~ 4 Mya through frequent hybridogenous speciation accompanied by the multiple evolution of partial endoreplication, a feature unique to orchids. The five most species-rich lineages radiated at an increased rate around the same time, which points to the existence of a common, probably environmental, trigger. In addition, the evolution of partial endoreplication strongly correlates with the diversification of clades possessing it. The full evolutionary consequences of partial endoreplication are yet to be understood, but species with PE save considerable resources by not performing full endoreplication and can invest them into producing larger flowers to better attract pollinators. Indeed, PE is more common in species with bigger flowers and larger genome sizes, so it might represent an effective evolutionary workaround reducing the cost of endoreplication in polyploids.

References

1. Givnish TJ *et al.* 2015 Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. Royal Soc. B* **282**, 20151553. (doi:10.1098/rspb.2015.1553)
2. Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, van den Berg C, Schuitman A. 2015 An updated classification of Orchidaceae. *Bot. J. Linn. Soc.* **177**, 151–174. (doi:10.1111/boj.12234)
3. Pérez-Escobar OA, Chomicki G, Condamine FL, Karremans AP, Bogarín D, Matzke NJ, Silvestro D, Antonelli A. 2017 Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytol.* **215**, 891–905. (doi:10.1111/nph.14629)
4. Leitch IJ, Dodsworth S. 2017 Endopolyploidy in plants. In *Encyclopedia of life sciences*, pp. 1–10. American Cancer Society. (doi:10.1002/9780470015902.a0020097.pub2)
5. Bory S, Catrice O, Brown S, Leitch IJ, Gigant R, Chiroleu F, Grisoni M, Duval M-F, Besse P. 2008 Natural polyploidy in *Vanilla planifolia* (Orchidaceae). *Genome* **51**, 816–826. (doi:10.1139/G08-068)
6. Brown SC *et al.* 2017 DNA remodeling by strict partial endoreplication in orchids, an

- original process in the plant kingdom. *Genome Biol. Evol.* **9**, 1051–1071.
(doi:10.1093/gbe/evx063)
7. Trávníček P, Čertner M, Ponert J, Chumová Z, Jersáková J, Suda J. 2019 Diversity in
genome size and GC content shows adaptive potential in orchids and is closely linked to
partial endoreplication, plant life-history traits and climatic conditions. *New Phytol.* **224**,
1642–1656. (doi:10.1111/nph.15996)
8. Trávníček P, Ponert J, Urfus T, Jersáková J, Vrána J, Hřibová E, Doležel J, Suda J. 2015
Challenges of flow-cytometric estimation of nuclear genome size in orchids, a plant
group with both whole-genome and progressively partial endoreplication. *Cytometry A*
87A, 958–966. (doi:10.1002/cyto.a.22681)
9. Bateman RM, Guy JJ, Rudall PJ, Leitch IJ, Pellicer J, Leitch AR. 2018 Evolutionary and
functional potential of ploidy increase within individual plants: somatic ploidy mapping
of the complex labellum of sexually deceptive bee orchids. *Ann. Bot.* **122**, 133–150.
(doi:10.1093/aob/mcy048)
10. Scholes DR, Paige KN. 2015 Plasticity in ploidy: a generalized response to stress. *Trends*
Plant Sci. **20**, 165–175. (doi:10.1016/j.tplants.2014.11.007)
11. Barow M. 2006 Endopolyploidy in seed plants. *BioEssays* **28**, 271–281.
(doi:10.1002/bies.20371)
12. Bhosale R *et al.* 2018 A spatiotemporal DNA endoploidy map of the *Arabidopsis* root
reveals roles for the endocycle in root development and stress adaptation. *Plant Cell* **30**,
2330–2351. (doi:10.1105/tpc.17.00983)
13. Karremans AP. 2016 Genera Pleurothallidinarum: an updated phylogenetic overview of
Pleurothallidinae. *Lankesteriana* **16**, 219–241. (doi:10.15517/lank.v16i2.26008)
14. Govaerts R *et al.* 2020 World checklist of Orchidaceae. Facilitated by the Royal Botanic
Gardens, Kew. Available from: <http://apps.kew.org/wcsp/> Retrieved 2020-02-01.
15. Karremans AP, Díaz-Morales M. 2019 The Pleurothallidinae: extremely high speciation
driven by pollinator adaptation. In *Proceedings of the 22nd World Orchid Conference*
(eds AM Pridgeon, AR Arosemena), pp. 363–388. Guayaquil, Ecuador: Asociación
Ecuatoriana de Orquideología.
16. Schmickl R, Liston A, Zeisek V, Oberlander K K, Weitemier K, Straub SCK, Cronn RC,
Suda J. 2016 Phylogenetic marker development for target enrichment from transcriptome
and genome skim data: the pipeline and its application in southern African *Oxalis*
(Oxalidaceae). *Mol. Ecol. Resour.* **16**, 1124–1135.
17. Karremans AP, Bogarín D, Díaz-Morales M, Fernández M, Osés L, Pupulin F. 2016

- 532 Phylogenetic reassessment of *Acianthera* (Orchidaceae: Pleurothallidinae). *Harv. Pap.*
533 *Bot.* **21**, 171–187. (doi:10.3100/hpib.v21iss2.2016.n4)
- 534 18. Bogarín D, Karremans AP, Fernández M. 2018 Genus-level taxonomical changes in the
535 *Lepanthes* affinity (Orchidaceae, Pleurothallidinae). *Phytotaxa* **340**, 128–136.
536 (doi:10.11646/phytotaxa.340.2.2)
- 537 19. Štorchová H, Hrdličková R, Chrtěk J, Tetera M, Fitze D, Fehrer J. 2000 An improved
538 method of DNA isolation from plants collected in the field and conserved in saturated
539 NaCl/CTAB solution. *Taxon* **49**, 79–84. (doi:10.2307/1223934)
- 540 20. Fér T, Schmickl RE. 2018 HybPhyloMaker: target enrichment data analysis from raw
541 reads to species trees. *Evol. Bioinform.* **14**, 1–9. (doi:10.1177/1176934317742613)
- 542 21. Borowiec ML. 2016 AMAS: a fast tool for alignment manipulation and computing of
543 summary statistics. *PeerJ* **4**, e1660. (doi:10.7717/peerj.1660)
- 544 22. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009 trimAl: a tool for automated
545 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973.
546 (doi:10.1093/bioinformatics/btp348)
- 547 23. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis
548 of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
- 549 24. R Core Team. 2018 R: a language and environment for statistical computing R
550 Foundation for statistical computing, Vienna, Austria. *R Foundation for Statistical*
551 *Computing, Vienna, Austria* <https://www.R-project.org/>
- 552 25. Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014
553 ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* **30**,
554 i541–i548. (doi:10.1093/bioinformatics/btu462)
- 555 26. Hodcroft E. 2016 *TreeCollapserCL4* (<http://emmahodcroft.com/TreeCollapseCL.html>).
556 See <http://emmahodcroft.com/TreeCollapseCL.html>.
- 557 27. Price MN, Dehal PS, Arkin AP. 2010 FastTree 2 – approximately maximum-likelihood
558 trees for large alignments. *PLoS One* **5**, e9490. (doi:10.1371/journal.pone.0009490)
- 559 28. Smith SA, Moore MJ, Brown JW, Yang Y. 2015 Analysis of phylogenomic datasets
560 reveals conflict, concordance, and gene duplications with examples from animals and
561 plants. *BMC Evol. Biol.* **15**, 150. (doi:10.1186/s12862-015-0423-0)
- 562 29. Rambaut A. 2018 *FigTree* v1.4.4. (<http://tree.bio.ed.ac.uk/software/>). See
563 <http://tree.bio.ed.ac.uk/software/>.
- 564 30. Huerta-Cepas J, Serra F, Bork P. 2016 ETE 3: reconstruction, analysis, and visualization
565 of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638. (doi:10.1093/molbev/msw046)

- 566 31. Poinar G, Rasmussen FN. 2017 Orchids from the past, with a new species in Baltic
567 amber. *Bot. J. Linn. Soc.* **183**, 327–333. (doi:10.1093/botlinnean/bow018)
- 568 32. Ramírez SR, Gravendeel B, Singer RB, Marshall CR, Pierce NE. 2007 Dating the origin
569 of the Orchidaceae from a fossil orchid with its pollinator. *Nature* **448**, 1042–1045.
570 (doi:10.1038/nature06039)
- 571 33. Drummond AJ, Bouckaert RR. 2015 *Bayesian Evolutionary Analysis with BEAST*.
572 Cambridge: Cambridge University Press. (doi:10.1017/CBO9781139095112)
- 573 34. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut
574 A, Drummond AJ. 2014 BEAST 2: A software platform for Bayesian evolutionary
575 analysis. *PLoS Comput. Biol.* **10**, e1003537. (doi:10.1371/journal.pcbi.1003537)
- 576 35. Åkerborg Ö, Sennblad B, Lagergren J. 2008 Birth-death prior on phylogeny and speed
577 dating. *BMC Evol. Biol.* **8**, 77. (doi:10.1186/1471-2148-8-77)
- 578 36. Battistuzzi FU, Billing-Ross P, Paliwal A, Kumar S. 2011 Fast and slow implementations
579 of relaxed-clock methods show similar patterns of accuracy in estimating divergence
580 times. *Mol. Biol. Evol.* **28**, 2439–2442. (doi:10.1093/molbev/msr100)
- 581 37. Ho SYW. 2014 The changing face of the molecular evolutionary clock. *Trends Ecol.*
582 *Evol.* **29**, 496–503. (doi:10.1016/j.tree.2014.07.004)
- 583 38. Mello B, Tao Q, Tamura K, Kumar S. 2017 Fast and accurate estimates of divergence
584 times from big data. *Mol. Biol. Evol.* **34**, 45–50. (doi:10.1093/molbev/msw247)
- 585 39. Kumar S, Hedges SB. 2016 Advances in time estimation methods for molecular data.
586 *Mol. Biol. Evol.* **33**, 863–869. (doi:10.1093/molbev/msw026)
- 587 40. Mello B. 2018 Estimating TimeTrees with MEGA and the TimeTree resource. *Mol. Biol.*
588 *Evol.* **35**, 2334–2342. (doi:10.1093/molbev/msy133)
- 589 41. Maddison WP, Midford PE, Otto SP. 2007 Estimating a binary character's effect on
590 speciation and extinction. *Syst. Biol.* **56**, 701–710. (doi:10.1080/10635150701607033)
- 591 42. FitzJohn RG. 2012 Diversitree: comparative phylogenetic analyses of diversification in
592 R. *Methods Ecol. Evol.* **3**, 1084–1092. (doi:10.1111/j.2041-210X.2012.00234.x)
- 593 43. Zeileis A, Hothorn T. 2002 Diagnostic Checking in Regression Relationships. *R News* **2**,
594 7–10.
- 595 44. Revell LJ. 2012 phytools: an R package for phylogenetic comparative biology (and other
596 things). *Methods Ecol. Evol.* **3**, 217–223. (doi:10.1111/j.2041-210X.2011.00169.x)
- 597 45. Bollback JP. 2006 SIMMAP: Stochastic character mapping of discrete traits on
598 phylogenies. *BMC Bioinform.* **7**, 88. (doi:10.1186/1471-2105-7-88)
- 599 46. Rabosky DL. 2014 Automatic detection of key innovations, rate shifts, and diversity-

- dependence on phylogenetic trees. *PLoS One* **9**, e89543.
(doi:10.1371/journal.pone.0089543)
47. Moore BR, Höhna S, May MR, Rannala B, Huelsenbeck JP. 2016 Critically evaluating
the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc.
Natl. Acad. Sci. U.S.A.* **113**, 9569–9574. (doi:10.1073/pnas.1518659113)
48. Rabosky DL, Mitchell JS, Chang J. 2017 Is BAMM flawed? Theoretical and practical
concerns in the analysis of multi-rate diversification models. *Syst. Biol.* **66**, 477–498.
(doi:10.1093/sysbio/syx037)
49. Title PO, Rabosky DL. 2019 Tip rates, phylogenies and diversification: What are we
estimating, and how good are the estimates? *Methods Ecol. Evol.* **10**, 821–834.
(doi:10.1111/2041-210X.13153)
50. Plummer M, Best N, Cowles K, Vines K. 2006 CODA: convergence diagnosis and
output analysis for MCMC. *R News* **6**, 7–11.
51. Rabosky DL, Grundler M, Anderson C, Title P, Shi JJ, Brown JW, Huang H, Larson JG.
2014 BAMMtools: an R package for the analysis of evolutionary dynamics on
phylogenetic trees. *Methods Ecol. Evol.* **5**, 701–707. (doi:10.1111/2041-210X.12199)
52. Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A,
Seehausen O. 2013 Genome-wide RAD sequence data provide unprecedented resolution
of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation.
Mol. Ecol. **22**, 787–798. (doi:10.1111/mec.12023)
53. Ramírez SR, Eltz T, Fujiwara MK, Gerlach G, Goldman-Huertas B, Tsutsui ND, Pierce
NE. 2011 Asynchronous diversification in a specialized plant-pollinator mutualism.
Science **333**, 1742–1746. (doi:10.1126/science.1209175)
54. Chomicki G, Bidel LPR, Ming F, Coiro M, Zhang X, Wang Y, Baissac Y,
Jay-Allemand C, Renner SS. 2015 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.
(doi:10.1111/nph.13106)
55. Givnish TJ *et al.* 2016 Orchid historical biogeography, diversification, Antarctica and the
paradox of orchid dispersal. *J. Biogeogr.* **43**, 1905–1916. (doi:10.1111/jbi.12854)
56. Shi JJ, Rabosky DL. 2015 Speciation dynamics during the global radiation of extant bats.
Evolution **69**, 1528–1545. (doi:10.1111/evo.12681)
57. Guo Y-Y, Luo Y-B, Liu Z-J, Wang X-Q. 2015 Reticulate evolution and sea-level
fluctuations together drove species diversification of slipper orchids (*Paphiopedilum*) in

- 634 South-East Asia. *Mol. Ecol.* **24**, 2838–2855. (doi:10.1111/mec.13189)
- 635 58. Brandrud MK, Baar J, Lorenzo MT, Athanasiadis A, Bateman RM, Chase MW, Hedrén
636 M, Paun O. 2020 Phylogenomic relationships of diploids and the origins of
637 allotetraploids in *Dactylorhiza* (Orchidaceae). *Syst. Biol.* **69**, 91–109.
638 (doi:10.1093/sysbio/syz035)
- 639 59. Russell A, Samuel R, Rupp B, Barfuss MHJ, Šafran M, Besendorfer V, Chase MW. 2010
640 Phylogenetics and cytology of a pantropical orchid genus *Polystachya* (Polystachyinae,
641 Vandaeae, Orchidaceae): Evidence from plastid DNA sequence data. *Taxon* **59**, 389–404.
642 (doi:10.1002/tax.592005)
- 643 60. Jones KE *et al.* 2019 An empirical assessment of a single family-wide hybrid capture
644 locus set at multiple evolutionary timescales in Asteraceae. *Appl. Plant Sci.* **7**, e11295.
645 (doi:10.1002/aps3.11295)
- 646 61. Crawl AA, Manos PS, McVay JD, Lemmon AR, Lemmon EM, Hipp AL. 2020
647 Uncovering the genomic signature of ancient introgression between white oak lineages
648 (*Quercus*). *New Phytol.* **226**, 1158–1170. (doi:10.1111/nph.15842)
- 649 62. Helmstetter AJ, Papadopoulos AST, Igea J, Van Dooren TJM, Leroi AM, Savolainen V.
650 2016 Viviparity stimulates diversification in an order of fish. *Nat. Commun.* **7**, 1–7.
651 (doi:10.1038/ncomms11271)
- 652 63. Davis MP, Midford PE, Maddison W. 2013 Exploring power and parameter estimation of
653 the BiSSE method for analyzing species diversification. *BMC Evol. Biol.* **13**, 38.
654 (doi:10.1186/1471-2148-13-38)
- 655 64. Rabosky DL, Goldberg EE. 2015 Model inadequacy and mistaken inferences of trait-
656 dependent speciation. *Syst. Biol.* **64**, 340–355. (doi:10.1093/sysbio/syu131)
- 657 65. Bhosale R, Maere S, De Veylder L. 2019 Endoreplication as a potential driver of cell
658 wall modifications. *Curr. Opin. Plant Biol.* **51**, 58–65. (doi:10.1016/j.pbi.2019.04.003)
- 659 66. Martel C, Cairampoma L, Stauffer FW, Ayasse M. 2016 *Telipogon peruvianus*
660 (Orchidaceae) flowers elicit pre-mating behaviour in *Eudejeania* (Tachinidae) males for
661 pollination. *PLoS One* **11**, e0165896. (doi:10.1371/journal.pone.0165896)
- 662 67. Millner HJ, Baldwin TC. 2016 Floral micromorphology of the genus *Restrepia*
663 (Orchidaceae) and the potential consequences for pollination. *Flora* **225**, 10–19.
664 (doi:10.1016/j.flora.2016.09.007)
- 665 68. Carotenuto G, Volpe V, Russo G, Politi M, Sciascia I, Almeida Engler J de, Genre A.
666 2019 Local endoreduplication as a feature of intracellular fungal accommodation in
667 arbuscular mycorrhizas. *New Phytol.* **223**, 430–446. (doi:10.1111/nph.15763)

69. Williamson B, Hadley G. 1969 DNA content of nuclei in orchid protocorms symbiotically infected with *Rhizoctonia*. *Nature* **222**, 582–583. (doi:10.1038/222582a0)
70. Rasmussen HN. 1990 Cell differentiation and mycorrhizal infection in *Dactylorhiza majalis* (Rchb. f.) Hunt & Summerh. (Orchidaceae) during germination *in vitro*. *New Phytol.* **116**, 137–147. (doi:10.1111/j.1469-8137.1990.tb00519.x)
71. Hynson NA, Madsen TP, Selosse M-A, Adam IKU, Ogura-Tsujita Y, Roy M, Gebauer G. 2013 The physiological ecology of mycoheterotrophy. In *Mycoheterotrophy: the biology of plants living on fungi* (ed V Merckx), pp. 297–342. New York, NY: Springer. (doi:10.1007/978-1-4614-5209-6_8)
72. Rasmussen HN, Dixon KW, Jersáková J, Těšitelová T. 2015 Germination and seedling establishment in orchids: a complex of requirements. *Ann. Bot.* **116**, 391–402. (doi:10.1093/aob/mcv087)

Acknowledgements

We are indebted to Blanka Hamplová and Štěpánka Hrdá from the BIOCEV centre for their help with sequencing and to F. Rooks for kindly improving the English.

Funding

The study was supported by the Czech Science Foundation (project no. 17-18080S) and the Ministry of Education, Youth and Sports of the Czech Republic (project NPU1: LO1417). It was also funded as part of a long-term research project of the Czech Academy of Sciences, Institute of Botany (RVO 67985939).

Author contributions

Z.C., E. Z., J.P. and P.-A.S. contributed equally to the work supervised by P.T. All authors carried out the analyses, interpreted and discussed the data and contributed to the writing and editing of the manuscript.

Data availability statement

Sequence data that support the findings of this study have been available in a public repository deposited in GenBank with the primary accession code PRJNA604559. The flow cytometric data that support the findings of this study are available from the corresponding author upon reasonable request.

Figure captions

Figure 1. Diversity tree for analyses of lineage diversification within the Pleurothallidinae based on a simplified species tree obtained using the ASTRAL approach. Clades are collapsed to 47 monophyletic lineages representing 40 recognized genera and coloured by extant species diversity. Clades with significantly increased net diversification rates are labelled by numbered and coloured asterisks. Clades are supplemented with pie charts reflecting the proportion of species based on their endoreplication type. Nodes are provided with pie charts reflecting an ancestral state reconstruction derived from stochastic character mapping of endoreplication type on the species tree.

Figure 2. Net diversification rates over time for the six clades with consistent and distinct shift configuration inferred from BAMM analysis. The graphs correspond to the clades marked in Fig. 1 and represent different types of endoreplication: (a) a clade of almost all Pleurothallidinae except two basalmost lineages, (b) a clade of the genus *Lepanthes* with the mixed type, (c) a clade of the genus *Dracula* with the partial type, (d) the genus *Masdevallia*, (e) a clade of the genus *Pleurothallis* with the conventional type, and (f) the genus *Stelis*. Coloured lines indicate net diversification rates and black lines represent net diversification rates of the background phylogeny. Shading around lines is for 90% confidence intervals.

Figure 3. State-dependent diversification rates extracted from BAMM analyses corresponding to three types of clades according to the type of endoreplication present – clades with exclusively conventional endoreplication (CE clades), clades with exclusively partial endoreplication (PE clades) and clades with both types (mixed clades). The graphs show the posterior distribution of rates for (a) net diversification, (b) speciation and (c) extinction, coloured by character state.

726 **Supporting information**

727 Electronic supplementary material is available online.

728

729 **Results S1.** Extended Results

730

731 **Methods S1.** Extended Materials and Methods

732

733 **Fig. S1.** Flow-cytometric histograms and scatter plots for two species of the genus *Lepanthes*
734 showing a different type of endoreplication.

735

736 **Fig. S2.** Ancestral state reconstructions of endoreplication type using stochastic character
737 mapping.

738

739 **Fig. S3.** Ancestral state reconstructions of endoreplication type using a full BiSSE model.

740

741 **Fig. S4.** Ancestral state reconstructions of endoreplication type using a BiSSE model with
742 equal extinction rates.

743

744 **Fig. S5.** Ancestral state reconstructions of endoreplication type using a BiSSE model with
745 equal speciation rates.

746

747 **Fig. S6.** Credible intervals of differences from the posterior distribution of samples taken
748 from BiSSE analyses.

749

750 **Fig. S7.** Species tree topologies obtained via the ASTRAL approach (a) and ML
751 concatenation (b) of Pleurothallidinae.

752

753 **Fig. S8.** Divergence time estimates and 95% credibility intervals (CIs) inferred for all nodes
754 of Pleurothallidinae chronograms.

755

756 **Fig. S9.** Mean phylorate plot of net diversification from BAMM analysis.

757

758 **Fig. S10.** The nine most common credible shift sets from BAMM analysis.

759

760 **Fig. S11.** Relative evidence of shifts in endoreplication mode inferred from the posterior
761 distribution of BAMM results.

762

763 **Fig. S12.** Credible intervals of differences from the posterior distribution of samples taken
764 from BAMM analyses and divided according to the type of endoreplication into partial (PE),
765 conventional (CE) and mixed clades.

766

767 **Fig. S13.** Posterior distribution of state-dependent rates extracted from BiSSE analyses with
768 equal extinction rates.

769

770 **Fig. S14.** State-dependent diversification rates extracted from BAMM analyses
771 corresponding to three types of clades according to the harboured type or types of
772 endoreplication – clades with exclusively conventional endoreplication (CE clades), clades
773 with exclusively partial endoreplication (PE clades) and clades with both types (mixed
774 clades).

775

776 **Table S1.** List of specimens included in the study supplemented by the type of
777 endoreplication, morphological traits and HybSeq statistics (as separate Excel table).

778

779 **Table S2.** List of genera according to the adopted taxonomic concept supplemented with the
780 number of species included, total species richness, endoreplication type, and data inferred
781 from ASTRAL topology.

782

783 **Table S3.** Summary characteristics for clades corresponding to genera with significantly
784 increased diversification rates.

Table 1. Analysis of variance (ANOVA) for selected morphological traits adopted from the literature with respect to the type of endoreplication (pairwise comparison of PE and CE taxa). P values both for ordinary and phylogenetic ANOVA are provided and values in bold are statistically significant ($\alpha = 0.05$). Arrows next to p-values indicate the direction of trait shifts for PE taxa.

Trait	F value	p _{ANOVA}	p _{phyloANOVA}
Ramicaul length	18.255	< 0.001 ↓	0.364
Leaf length	2.724	0.010 ↑	0.717
Leaf width	0.079	0.779	0.954
Leaf size (length x width)	0.186	0.667	0.937
Sepal length	87.339	< 0.001 ↑	0.038
Petal length	0.7	0.404	0.863
Sepal length / leaf size	13.5	< 0.001 ↑	0.457
Sepal length / ramicaul length	62.366	< 0.001 ↑	0.08
Sepal length / ramicaul + leaf length	84.558	< 0.001 ↑	0.037

(a)

(b)

(c)