Article type: Original Article

Running head: Andean effect on the swan orchids biogeography

THE IMPROBABLE JOURNEYS OF EPIPHYTIC PLANTS ACROSS THE ANDES: HISTORICAL BIOGEOGRAPHY OF CYCNOCHES (CATASETINAE, ORCHIDACEAE)

Oscar Alejandro Pérez-Escobar1*,†, Marc Gottschling1, Guillaume Chomicki1, Fabien L. Condamine2, Bente Klitgård3, Emerson Pansarin4 and Günter Gerlach5*

1 Department Biologie, Systematische Botanik und Mykologie, GeoBio-Center, Ludwig-Maximilians Universität, Menzinger Straße 67, D - 80638, Munich, Germany
2 CNRS, UMR 5554 Institut de Sciences de l'Evolution (Université de Montpellier), Place Eugène Bataillon, 34095 Montpellier, France
3 Identification and Naming department, Royal Botanic Gardens, Kew, TW9 3AB, Surrey, UK.
4 Departamento de Biologia, Facultade de Filosofia, Ciências e Letras, Universidade de Sao Paulo, Ribeirao Preto, SP, 14040-901, Brazil
5 Botanischer Garten München, Menzinger Straße 61, D - 80638, Munich, Germany

†Current address: Identification and Naming department, Royal Botanic Gardens, Kew, TW9 3AB, Surrey, UK
*Corresponding author: gerlach@extern.lrz-muenchen.de; o.perez-escobar@kew.org

Number of words: 4101 (main text) and 7052 (including references and captions)
Number of figures: 4
Number of tables: 2
Abstract [200 words]:

The Andean uplift is one of the major orographic events in the New World and has impacted considerably the diversification of numerous Neotropical organisms. Despite its importance for biogeography, the specific role of mountain ranges as a dispersal barrier between South and Central American lowland plant lineages is still poorly understood. The swan orchids (*Cycnoches*) comprise ca 34 epiphytic species distributed in lowland and pre-montane forests of Central and South America. Here, we study the historical biogeography of *Cycnoches* to better understand the impact of the Andean uplift on the diversification of Neotropical lowland plant lineages. Using novel molecular sequences (five nuclear and plastid regions) and twelve biogeographic models with and without founder-event speciation, we infer that the most recent common ancestor of *Cycnoches* may have originated in Amazonia ca 5 Mya. The first colonization of Central America occurred from a direct migration event from Amazonia, and multiple bidirectional trans-Andean migrations between Amazonia and Central America took place subsequently. Notably, such biological exchange occurred well after major mountain building periods. The Andes have not acted as an impassable barrier for epiphytic lowland lineages such as orchids having a great potential for effortless dispersal because of the very light, anemochorous seeds.

Key words: ancestral area, Andean uplift, barrier, dispersal, historical biogeography, molecular clock.
Introduction

Neotropical landscape and biodiversity have long drawn the attention of naturalists. The tropical Andes are of particular interest as the world’s premier biodiversity hotspot, with both an extraordinary species richness and a remarkable level of endemism. The combination of molecular phylogenies with species distributions and the fossil record has uncovered different biotic and abiotic factors that fostered diversification in the Neotropics. However, biogeographical studies applying modern phylogenetic methods are only available for few Neotropical plant clades (e.g., Begonia, Cyathostegia, Heliotropium, Lupinus, palms, Rubiaceae). These studies have generally demonstrated the importance of geological processes such as mountain building and establishing the Isthmus of Panama for the diversification of Neotropical plants.

One of the most biologically relevant abiotic processes in the diverse geological history of the Americas is the rise of the Andes. Andean mountain building was driven by plate tectonic re-adjustments that started during the Palaeogene and continued until the Pliocene. The fossil record (e.g., palynological and geological data: isotope measurements, sediment loads, apatite fission-track data) collectively indicate that the Andean uplift was a partially constant process punctuated by periods of intensified mountain building. Two of the most intense uplift periods occurred around 12 (mid-Miocene) and 4.5 million years ago (Mya; early Pliocene). During these periods, the Northern Andes reached elevations as high as 4500 m in the Pliocene, whereas the Central Andes already peaked an altitude of 4500 m during the mid-Miocene.

Newly formed mountain ranges may had a strong impact on the adjacent Amazonian landscapes and the inhabiting organisms due to the transformation of its
drainage systems. The Andes also influenced local and regional climates by forming the only barrier to atmospheric circulation in the Neotropics. The rise of the Andes led to the formation of island-like habitats and of local microclimates and soil conditions that eventually fostered species diversification. At the same time, the Andes provided physical and/or ecological barriers to species dispersal and migration. For instance, the efficiency of the Northern Andes as migration barrier is shown for Andean centred woody species of Crematosperma, Klarobelia, Malmea, and Mosannona (Annonaceae). Propagules are dispersed by animals in these lineages, and none of their constituent species occur in both east and west of the Andes mountain range.

Recent phylogenetic studies provide solid evidence for the important role of Andean uplift in diversification of highland-dwelling plant groups (e.g., Lupinus, Bartsia, centropogonid Campanulaceae). However, the impact of such orographic processes for the lowland flora is still poorly understood. Thus, the question remains whether Andean uplift has indeed been an abiotic barrier to migration for epiphytic lineages such as lowland orchids and bromeliads, both being important components of Neotropical forests.

Epiphytic diversity is dramatically greater in the Neotropics than in any other tropical region around the world, being twice as high than, for instance, in Australasia. Several traits shared by Neotropical epiphytic taxa, and related to their reproductive biology, may explain such overwhelming differences in diversity. One of the most prominent shared traits is the lightness and the very small size of the propagules, occasionally with highly elaborated surfaces (e.g., bromeliads, ferns, orchids, Pinguicula). The potential of dust-like seeds for longer distance dispersals by wind might be much greater compared to other plants with propagules locally dispersed.
by animals (e.g., Araceae). Nonetheless, it remains largely unknown whether lowland epiphyte lineages have been able to disperse across geological barriers such as the Andes. Well-sampled phylogenies for epiphyte clades have been lacking to address this issue.

Several wind-dispersed plant lineages (e.g., *Begonia*; bromeliads) span across the Neotropical region, many of which are restricted to lowland elevations. One example is the orchid tribe Cymbidieae comprising ca 3000 species that are mostly distributed in the Neotropics (but with few representatives in the Old World Tropics).

To Neotropical Cymbidieae belong the swan orchids (*Cycnoches*) that are known for the striking sexual dimorphisms and pollination syndrome. Molecular phylogenetics and morphological studies conducted to date confirm the inclusion of *Cycnoches* within Catasetinae and its sister group relationship to *Mormodes*.

*Cycnoches* encompasses 34 species of epiphytes that are distributed from Southern Mexico to Central Brazil and Bolivia. Its species are best represented in the Amazonian forests of Brazil, Ecuador, and Peru, but also in the Llanos and Caribbean regions of Colombia and Venezuela. They commonly inhabit lowland tropical wet forests, ranging from 0 to 800 m. *Cycnoches* species are pollinated by male euglossine bees, which collect volatile compounds produced by flowers but also from other sources (e.g., rotten wood, faeces). They have rather a restricted geographical range, and most of them are distributed in single biogeographical area. Nevertheless, one species (*C. chlorochilon*) is distributed in both sides of the Northern Andes.

Because of the striking sexual system they have evolved, swan orchids have long attracted the attention of several prominent botanists including. Despite this long interest, previous phylogenetic studies have only been included up to three species of *Cycnoches* probably because of their scarcity both in the wild and in herbarium.
collections\textsuperscript{42,43,45,46,50}. Moreover, the lack of a solid and well-sampled phylogeny of \textit{Cycnoches} has precluded addressing specific questions concerning the role of Andean uplift in their biogeographical history. The narrow geographical distribution of almost all extant \textit{Cycnoches} species and their restricted habitat preference given (i.e., lowland wet forests, see above), we expect the swan orchids diversification to be strongly influenced by the Andean uplift. In particular, we hypothesize that Andean uplift was an impassable, isolative, barrier for \textit{Cycnoches}, as already reported for other plant lineages. To test this hypothesis, we generated a strongly supported 5-loci phylogeny, sampling 23 out of 34 accepted \textit{Cycnoches} species and comprising its known diversity and distribution, and performed models of biogeographical analyses.

**Results**

**Phylogeny of Cycnoches**

Our phylogeny comprised 23 out of 34 accepted species of \textit{Cycnoches}. Table S5 provides detailed alignment descriptions. The concatenated ‘nuc’ alignment was 2395 bp in length and included 310 parsimony informative sites, and the concatenated ‘cp’ alignment was 2419 bp and comprised 171 parsimony informative positions. Independently derived ‘nuc’ and ‘cp’ phylogenies revealed topologies with conflicting and highly supported phylogenetic placements. PACo analysis revealed 22 potential outlier OTUs (see below; Figs S1–S2) belonging to \textit{Cycnoches} (majority of 19), \textit{Dressleria}, and \textit{Mormodes}. After inspection of the potentially conflicting OTUs placement in ‘nuc’ and ‘cp’ phylogenies (Fig. S2), 20 outliers were confirmed as conflicting terminals. Thus, only two terminals (i.e., \textit{Dressleria severiniana} and one representative of \textit{C. lehmannii}) were misclassified by PACo as conflicting (see Appendix S1 for a detailed explanation on outlier handling).
Within *Cycnoches*, the ‘nuc’ phylogeny recovered three maximally supported clades (A, B, C; Fig. 1). Clade A included all sequenced accessions of *Cycnoches hagii* and was recovered as sister group of the remaining species of *Cycnoches* comprising clades B and C. By contrast, the ‘cp’ phylogeny showed two main maximally supported clades (namely I and II), each including a set of intermingled species from clades A, B, and C. The best ML tree inferred from non-conflicting, concatenated ‘nuc’ and ‘cp’ datasets showing the internal phylogenetic relationships of *Cycnoches* is presented in Figure 2. Virtually, all internal nodes of the backbone phylogeny were highly, if not maximally supported by MLBS and PP values. *Cycnoches* segregated into three main lineages (clades A, B and C). Clade A (i.e., all specimens of *C. haagii*) was sister group of the remaining species of *Cycnoches* clustering in clades B and C.

**Molecular dating of Cycnoches**

Comparison of marginal likelihood estimates (MLE) of tree priors and clock models revealed that an uncorrelated molecular clock combined to a birth-death tree prior with incomplete sampling best fitted our data (MLE of –14244.47; Table 1). Nevertheless, Bayes factors (BF) clearly rejected the strict clock models (BF = 72, 45, and 40 for all tree priors), but did not provide strong evidence to support an incomplete sampling birth-death model vs a standard birth-death model (BF=1.08; Table 1). Analysis of the log file produced by dating analyses under the relaxed clock and tree models are also shown in Table 1. Overall, they yielded CV values between 0.35 and 0.42 (indicating there was among branch rate heterogeneity, which argued for the use of a relaxed molecular clock). Therefore, results obtained from a dated phylogeny with relaxed molecular clock and birth-death standard speciation model (Table 1) are presented only (Figs 3, S5).
A chronogram showing absolute ages estimated under a relaxed clock is presented in Figure 3 (see also Fig. S5 for the 95% confidence intervals) and shows that *Cycnoches* and *Mormodes* shared a common ancestor during the beginning of the late Miocene (9.1 Mya ±3). Diversification of *Cycnoches* took place around 5 Mya ±2 at end of the late Miocene. The split between *Cycnoches* clades B and C occurred during late Pliocene (3 Mya ±2), whereas both clades were estimated to be of Pleistocene ages (1.2 and 1.6 Mya ±1, respectively).

**Ancestral area estimations**

Table 2 provides ML statistics for the biogeographical models as inferred in BioGeoBEARS. The best fitting parametric model was the BAYAREA star, including the founder-event speciation. This model revealed a geographical origin in Amazonia as the most likely ancestral area of *Cycnoches* (Figs 3, S6). The MRCA of clades B and C was reconstructed to have inhabited in Amazonia region. The MRCA of clade C occurred in Central America, whereas the MRCA of clade B inhabited in Amazonia region.

Three independent trans-Andean migration events between Amazonia and Central America could be identified. The first migration from Amazonia to Central America took place towards the late Pliocene (±1 Mya), after the divergence of MRCA of clades B and C. A second migration from Central America to Amazonia took place around mid-Pleistocene (±1 Mya) by the MRCA of *C. maculatum, C. manoelae, C. peruvianum, C. quatuorcrisitis*, and *C. suarezii*. Last biotical exchange from Amazonia to Central America was dated to late Pleistocene (±0.5 Mya) with the MRCA of *C. chlorchilon, C. ventricosum*, and *C. warzsewiczii*. 
Discussion

Influence of Andean orogeny on the biogeography of a Neotropical epiphyte group

Our study provides a solid phylogenetic framework, divergence times, and ancestral areas estimation for *Cycnoches*. Central America has been considered the most likely region of *Cycnoches* origin, probably because of its locally elevated species richness as compared to other areas in the Neotropics (Romero & Gerlach, in press). However, this evolutionary scenario is rejected by our analyses, which instead support a South American origin of *Cycnoches* in the late Miocene (ca. 5 Mya ±2; Fig. 3, Fig. S6). The ancestral area estimated at *Cycnoches* MRCA largely reflects the current distribution of early diverging lineages such as *C. haagii*, a species today distributed in Amazonia and the Guiana Shield. Thus, the early diversification of *Cycnoches* has taken place well after the most intense mountain building events of the Northern and Central Andes ca. 12 and 10 Mya, respectively. Our results indicate a less important role of the Andes as a biogeographic barrier for the diversification of *Cycnoches*. Two migrations from Amazonia to Central America and one reverse colonisation event back imply vigorous ancient dispersals across the Andes during last 5 million years (Figs 3). During the early Pliocene (i.e., the period when *Cycnoches* has started to diversify), the Northern Andes of Colombia and Venezuela have already reached elevations up to at least 3000 m (Hoorn et al., 2010). Moreover, three migrations from Amazonia to Central America and back, respectively, took place ~2 Mya, when the Northern and Central Andes already peaked at a mean elevation of 4000 m (see mean Andean elevation displayed in the inset of Fig. 3). Similar trans-Andean migrations are reported for bromeliads and ferns. In Bromeliaceae (subfamily Bromelioidae), direct migrations from the Brazilian Shield towards Central America took place around 7 Mya, a period where Northern Andes
reached a palaeoelevation of ~2000 m\(^7\). Furthermore, biotical exchanges between South and Central America have also been reported in the fern complex *Jamesonia-Eriosorus* (Pteridaceae)\(^30\). Here, migrants from the Brazilian coast colonised and further established in Central Andes, and from there subsequently migrated towards Central America during the late Pleistocene\(^55\).

Impermeability of the Andes as a barrier for reproductive isolation does not appear as important for Neotropical epiphyte lineages such as *Cycnoches*. Our results also point to a rapid colonization of the mainland Neotropics (i.e., colonization of Central America, Amazonia, and Choco in ~5 million yr.). Such biotic invasion may be related to the biology of the group having anemochorous seed dispersal and an epiphytic habit. Most orchid seeds are characterized by their minute size and reduced weight, as well as by their elaborated seed coats\(^56,57\). These traits allow them to easily remain airborne for extended periods of time and to travel over long distances\(^57\). *Cycnoches* species might have good dispersal potential as well because of their seeds, which are between 100–300 µm long, 50–60 µm wide, and about 3.6 µg weight\(^56,58\).

Empirical evidence strongly supports short mean dispersal distances (e.g., 4–5 m) of seeds from terrestrial orchids such as *Orchis purpurea*\(^59\), yet there are no experimental studies supporting putatively longer dispersal of epiphytic orchid seeds. Nevertheless, disjunctive distributions of several orchid taxa occurring, for instance, in remote islands (e.g., *Anacamptis pyramidalis* growing in Aldabra island and Madagascar\(^60\)) require dispersals over longer ranges as explanation. Authors have already reported disjunctive distributions of several orchid species resulting from putative long-distance dispersals\(^61,62\). However, to the best of our knowledge, no study has yet reported ancient orchid long-distance dispersal across a geographical barrier such as the Andes.
Phylogenetic conflict between nuclear and plastid phylogenies in *Cycnoches*

Our study brings important insights for the species relationships within *Cycnoches*. Previous phylogenetic studies about Catasetinae have included not more than three species of *Cycnoches* \(^{42,43,45,46,50}\), hence keeping internal relationships (and corresponding conflicts, see below) of the lineages unresolved. Serious phylogenetic incongruence between ‘nuc’ and ‘cp’ tree topologies of Catasetinae has been firstly identified by \(^{63}\), but a discussion about the respective plausibility of the trees has been remained undone.

All major clades in our ‘nuc’ phylogeny are consistent with morphological concepts of *Cycnoches* (Fig. 1a). *Cycnoches hagii* (Clade A) differs from other *Cycnoches* species by the fleshy oblong, curved pair set of calli located towards the base of the labellum. Clade B includes species with conspicuously large male flowers and an entire to 4-lobed labellum blade. Clade C comprises species with proportionately small male flowers and labella with 8–10 marginal projections \(^{48}\). The consistence between ‘nuc’ molecules and morphology is challenged by an apparent correlation between the ‘cp’ molecular tree and the distributions of particular lineages (Fig. 1b): Clade I comprises species occurring mostly in Amazonia, Caribbean-Llanos, and Guyana shield regions (except *C. warzsewiczii* present in Panama and Costa Rica: Fig. S4). Clade II, in turn, includes species exclusively distributed in Central America and Choco regions.

Topological incongruence between phylogenies derived from different DNA data partitions is a widespread phenomenon in phylogenetic inferences \(^{64,65}\). Examples from angiosperms (e.g., Araceae \(^{37}\), Asteraceae \(^{66}\), Saxifragaceae \(^{67}\)) have long revealed similar conflicting patterns, in which the ‘nuc’ phylogeny is in accordance with morphology, whereas the ‘cp’ relationships correlate to geographical distributions. For
the particular case of Cycnoches, hybridization might have explanatory power for the nuclear-plastid conflict observed in the clade. Euglossine-bee-pollinated orchids such as Cycnoches produce a blend of volatile compounds, which attracts male Euglossine bees, and pollination takes place while bees collect such compounds produced by specialized tissues in the flower. Species-specific production of floral blends and therefore attraction of a unique set of pollinator(s) has been accounted as an isolative reproductive barrier in Euglossine bee pollinated orchids. Nevertheless, intra-specific variation of fragrances produced by the flower has been reported in several orchid lineages such as Stanhopea and even Cycnoches. Fragrance variation may result on attraction of a set of pollinators that are shared by species co-occurring in the same biogeographical region (e.g. C. dianae, C. guttulatum, and C. pachydactylon; see also Fig. S4) with similar composition of the fragrance profile, providing an opportunity to hybridization to occur. Little is known about pollinators of Cycnoches, but our own observations of pollinator sharing between species and presence of polymorphic species may indicate past hybridization processes.

**Conclusion**

Based on a solid, comprehensively sampled phylogeny we provide macroevolutionary evidence for a South American origin of Cycnoches. Our biogeographical analysis indicates colonization of Central America via a direct migration from the Amazonian basin. More importantly, the analyses support three recent trans-Andean, bidirectional migration processes between Central America and Amazonia, which is indicative for a minor effect of the Andean barrier on swan orchids migration. Consequently, our study enlightens the limited role of Andean mountain
building on the range evolution and diversification of lowland Neotropical epiphytic lineages.

**Material and methods**

**Taxon sampling, DNA sequencing and phylogenetic analyses**

Species names, geographical origins, voucher specimens, and GenBank accession numbers of sequences included in phylogenetic analyses are provided in Table S1. Our study builds-up upon the DNA data matrices generated by\textsuperscript{40,46,63}. Genomic DNA was extracted from herbarium and fresh leaf material with the NucleoSpin® plant kit (Macherey-Nagel; Düren, Germany). We amplified and sequenced nuclear (consistently referred as ‘nuc’ henceforth) ribosomal external and internal transcribed spacers (ETS and ITS, respectively), and a fragment of the \textit{Xdh} gene. We also sequenced a ~1500 bp fragment of the plastid (referred to as ‘cp’) gene \textit{ycf1}, as well as the \textit{trnS–trnG} intergenic spacer. Amplification and sequencing settings, as well as sequencing primers used for ITS, ETS, \textit{Xdh}, \textit{trnS–trnG}, and \textit{ycf1} are the same as in\textsuperscript{63,72} (Table S2). In this study, 84 sequences were newly generated (Table S1). Loci were aligned separately using MAFFT 7.\textsuperscript{73}. For ‘nuc’ ribosomal RNA loci and ‘cp’ \textit{trnS–trnG} spacer, secondary structure of molecules was taken into account (i.e., the -qINSi option). Congruence between ‘nuc’ and ‘cp’ datasets was assessed following Pérez-Escobar et al.\textsuperscript{63} and using PACo\textsuperscript{74}. The procedure is available as a pipeline (http://www.uv.es/cophylpaco/) and was also employed to identify operational terminal units (OTUs) from the ‘cp’ dataset that conflicting with the ‘nuc’ dataset (potential outliers detected by PACo are shown in Figs S1–S2). A detailed explanation on PACo and a rationale on outlier handling is provided Extended Materials and Methods of Appendix S1.
Phylogenetic analyses of separate and concatenated loci were carried out under maximum likelihood (ML) and Bayesian inference (BI). The best-fitting evolutionary models for ML and Bayesian analyses (for each data partition) were selected using jModelTest v.2.1.6, relying on a Likelihood Ratio Test (LRT) and the Akaike information criterion (AIC) (Table S3). Phylogenetic inference relied on the ML approach implemented in RAXML-HPC v.8.2.4 and BI as implemented in MRBAYES v.3.2.2 and were carried out on the CIPRES Science Gateway computing facility. Bayesian inferences were performed with two independent runs, each with four Markov chains Monte Carlo (MCMC) running for 30 million generations each, and sampled every 1000th generation (all other prior settings by default). Log files derived from MRBAYES were examined, and the convergence of MCMC was assessed using TRACER (available at: http://tree.bio.ed.ac.uk/software/tracer/). Node support values were assessed for both the ML tree using Maximum Likelihood Bootstrap Support (MLBS) and the consensus Bayesian tree using posterior probabilities (PP).

Molecular clock dating

A few orchid macrofossils are available for Orchidaceae, but these are assigned to lineages very distantly related to our groups of interest. Using distant outgroups to calibrate our Cycnoches phylogeny would have created extensive sampling heterogeneities, which can result in spurious results. Thus, we had to rely on secondary calibrations. In order to obtain the best secondary calibration points possible, we first generated an Orchidaceae-wide fossil-calibrated phylogenies, sampling orchid species sampled as evenly as possible along the tree. Loci, number of sequences and settings for absolute age estimation of the Orchidaceae-wide fossil calibrate phylogeny are provided in the Extended Materials and Methods of Appendix S1. The
ages obtained were very similar to recent orchid dating studies\textsuperscript{82,83} and the dated phylogeny is shown in Figure S3.

We selected two secondary calibrations for the dating of \textit{Cycnoches}: (i) the crown group of Catasetinae was set to 19.8 Mya with a standard deviation of 4 to reflect the 95\% CI, and (ii) and the root of the \textit{Cycnoches} tree (i.e., MRCA of \textit{Cyrtopodium} + Catasetinae) was set to 27.1 Mya with a standard deviation of 6. To explore the clock-likeness of the data, we used both strict clock and uncorrelated lognormal clock models, and compared different tree priors (pure-birth, standard birth-death, and incomplete sampling birth-death). For strict molecular clock calibration, we placed a single constraint only at the tree root (27.1 Mya with a standard deviation of 6) using a normal distribution. The best-fitting tree speciation model was selected using Bayes factors calculated from marginal likelihoods computed for every model using the stepping-stone sampling\textsuperscript{84} (Table 1). For each clock model, we ran two MCMC analyses with 20 million generations each, sampled every 1000\textsuperscript{th} generation. For the relaxed molecular clock analyses, we estimated the coefficient of variation (CV) to inform us on the rate heterogeneity among branches (CV approaching 0 indicates that a strict clock model cannot be rejected). Parameter convergence was confirmed using TRACER (http://tree.bio.ed.ac.uk/software/tracer/). All dating analyses were performed at the CIPRES Science Gateway computing facility\textsuperscript{78}.  

\textit{Ancestral area estimations}

Species ranges of \textit{Cycnoches} of were coded from the collection site and type locality of the material sequenced, which reflect the distribution ranges for every taxa included in our phylogeny (except by \textit{C. chlorochilon} and \textit{C. pentadactylon}, which also occur in Central America and Southeastern South America, respectively; see bellow).
Distribution data of *Cycnoches* and outgroup taxa was obtained from own field observations, literature\(^{85,86}\) and from herbarium specimens (Herbaria AMES, COL, F, M, MO, SEL, US); this information was employed to code outgroup distribution ranges. Biogeographical areas were derived from distribution maps of the orchids under investigation (Fig. S4) as well as species distributions observed in other plant lineages (e.g., Rubiaceae\(^6\); Bromeliaceae\(^38\)). We coded for eight biogeographical areas using the R-package ‘SpeciesGeocodeR’\(^{87}\): (1) Central America comprising southern Mexico through Panama; (2) Caribbean-Llanos comprising the coastal northernmost areas and plains of Colombia and Venezuela\(^{88}\); (3) Guiana Shield encompasses areas above 800 m in Colombia, Venezuela, Brazil, Guyana, Suriname, and French Guiana; (4) Amazonia encompassing lowlands and pre-montane forest below 800 m in Colombia, Ecuador, Peru, Brazil, Venezuela, Guyana, Suriname, and French Guiana\(^6\); (5) Chocó comprising lowlands below 500 m of the western Andes in Colombia and Ecuador; (6) Northern Andes including elevated areas above 800 m from Southernmost Peru to Northern Colombia and Northeast Venezuela; (7) Central Andes comprising areas above 800 m in Northern Peru to Northern Chile and Northeast Argentina; (8) South-eastern South America encompassing part of the Brazilian shield, the Atlantic forest, South-eastern Bolivia, Paraguay, Uruguay, and Northern Argentina. A map with the eight biogeographical areas, number of species for each area, and the corresponding distribution records of *Cycnoches* plus outgroups are provided in Figure S4.

To infer ancestral areas in *Cycnoches*, we used the R-package ‘BioGeoBEARS’ (Biogeography with Bayesian and Likelihood Evolutionary Analysis in R script\(^{89}\)). Using BioGeoBEARS, we tested the fit of six biogeographic models with and without founder-event speciation (or jump speciation), altogether testing the role and contribution of evolutionary processes that were taken into account to explain today’s
observed distributions (i.e., range expansions, local extinctions, founder-event speciation, vicariance, and speciation despite sympatry) in a joint statistical framework.

It is therefore capable of model testing and of determining, which process fits better the geographical and phylogenetic data for any particular clade.


29. Uribe-Convers, S. & Tank, D. C. Shifts in diversification rates linked to biogeographic movement into new areas: An example of a recent radiation in the


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


529 51. Romero, G. A. & Gerlach, G. in Flora Mesoamericana (Missouri Botanical
530 Garden Press).

531 52. Garzione, C. N., Molnar, P., Libarkin, J. C. & Macfadden, B. J. Rapid late
532 Miocene rise of the Bolivian Altiplano: evidence for removal of mantle


535 54. Garzione, C. N. et al. Clumped isotope evidence for diachronous surface cooling
536 of the Altiplano and pulsed surface uplift of the Central Andes. Earth Planet. Sci.

538 55. Sánchez-Baracaldo, P. & Thomas, G. H. Adaptation and convergent evolution
539 within the Jamesonia-Eriosorus complex in high-elevation biodiverse andean

541 56. Arditti, J. & Ghani, A. K. A. Numerical and physical properties of orchid seeds

543 57. Jersáková, J. & Malinová, T. Spatial aspects of seed dispersal and seedling

545 58. Barthlott, W., Große-Veldmann, B. & Korotkova, N. Orchid seed diversity: a
546 scanning electron microscopy survey. (Botanic Garden and Botanical Museum
547 Berlin, 2014).

548 59. Jacquemyn, H. et al. A spatially explicit analysis of seedling recruitment in the

550 60. Close, R. C., Moar, N. T., Tomlinson, A. I. & Lowe, A. D. Aerial dispersal of
551 biological material from Australia to New Zealand. Int. J. Biometeorol. 22, 1–19
552 (1978).


81. Drummond, A. J. & Bouckaert, R. R. *Bayesian evolutionary analysis with


Acknowledgements

We are grateful to Martina Silber (Munich, Germany) for valuable assistance in the lab, to Mario Blanco (San José, Costa Rica), Diego Bogarín (San José, Costa Rica), Norman Cash and family (Managua, Nicaragua) for the great help provided during field trip, to Boris Schlumpberger (Hannover, Germany), Andres Maduro (Panama City, Panama), Ivan G. Alvarez (David, Panama), Lankester Botanical Garden (Cartago, Costa Rica), Gustavo Romero (Herbarium HUH, Cambridge, U.S.A) and University of Costa Rica for providing plant material. We also thank the Smithsonian Tropical Research Institute (Panama City) and Costa Rican Ministry of Environment (MINAET) for providing research permits (No. SE/AP-20-13 and 026-2013-SINAC, respectively). We finally thank Alexandre Antonelli (Gothenburg, Sweden), for granting access to graphical material employed in Fig. 3. This work was supported by the Colombian Administrative Department of Science and Technology (COLCIENCIAS; granted to OAPE). We thank the Deutsche Forschungsgemeinschaft (granted to GG, GE 828/12-1) for financial support.

Author contributions

O.A.P.E., M.G. and F.L.C. designed research; G.G. and O.A.P.E. collected samples; O.A.P.E. and E.P. performed all the lab work; O.A.P.E., F.L.C., and G.C. performed all analyses; O.A.P.E., M.G., F.C., G.C., G.G., B.K. and E.P. wrote the manuscript under the lead of O.A.P.E., M.G, and F.C.
653  **Competing financial interest**

654  The authors declare no competing financial interest.
Tables

Table 1. Marginal likelihoods of tree speciation models for relaxed and strict molecular clocks.

Table 2. Likelihood of biogeographic models as implemented in DEC, DIVA-like and BayArea-like.

Figures

Figure 1. Best scoring, ML tree of Cycnoches obtained from concatenated A) nuclear ETS, ITS, Xdh and B) plastid trnS-G, ycf1 loci. Node charts indicate Bootstrap Support (BS > 75), in where fully red diagrams indicate BS of 100. Numbers at nodes indicate Bayesian Posterior Probability (PP > 0.95). Representatives of each clade are shown in pictures: 1) C. haagii; 2) C. ventricosum (left) and C. pentadactylon (right); 3) C. herrenhusanum (left) and C. peruvianum (right). The red arrows indicate autapomorphies of every clade: in 1) fleshy oblong, curved pair set of calli located towards the base of the labellum, 2) labellum blade entire to 4-lobed, and 3) labellum with 8-10 marginal projections. Geopolitical boundaries map generated by ArcMAP software (http://www.esri.com), using layers freely available at http://www.diva-gis.org/Data. Photos: O. Pérez & G.Gerlach.

Figure 2. Best scoring, ML tree of Cycnoches obtained from concatenated, non-conflicting ‘nuc’ ETS, ITS, and Xdh and ‘cp’ trnS–trnG and ycf1 loci. Node charts indicate Bootstrap Support (BS > 75), in where fully red diagrams indicate BS of 100. Numbers at nodes indicate Bayesian Posterior Probability (PP > 0.95).

Figure 3. Chronogram for Cycnoches obtained under a relaxed clock model, applied to a non-conflicting, concatenated ‘nuc’ (ITS, ETS, Xdh) and ‘cp’ loci (trnS–trnG, ycf1). Age
estimates, including maximum and minimum intervals for all nodes, are provided in Figure S5. Time scale is provided in million years ago (Mya). Node charts correspond to ancestral areas estimated under the BayArea-Like* model, including founder event process (J). Pink, orange and blue lines indicate mean elevations (m) on Colombian and Venezuelan Northern Andes, respectively (adapted from Hoorn et al., 2010). (Inset) Coded areas used for biogeographical analysis. Geopolitical boundaries map generated by ArcMAP software (http://www.esri.com). Political divisions and elevation data from DIVA-GIS (http://www.diva-gis.org/gdata).

**Supplementary material**

**Appendix S1.** Extended Materials and Methods, supplementary tables and figures.
### Table 1

<table>
<thead>
<tr>
<th>Tree prior</th>
<th>MLE (SS)</th>
<th>CV</th>
<th>Yule</th>
<th>BD</th>
<th>BD sampling</th>
<th>Yule</th>
<th>BD</th>
<th>BD sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strict clock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yule</td>
<td>-14280.93</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>BD</td>
<td>-14267.36</td>
<td>-</td>
<td>27.14</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BD sampling</td>
<td>-14264.73</td>
<td>-</td>
<td>32.4</td>
<td>5.26</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Relaxed clock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yule</td>
<td>-14259.53</td>
<td>0.42</td>
<td>42.8</td>
<td>15.66</td>
<td>10.4</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BD</td>
<td>-14245.01</td>
<td>0.35</td>
<td>71.84</td>
<td>44.7</td>
<td>39.44</td>
<td>29.04</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>BD sampling</td>
<td>-14244.47</td>
<td>0.35</td>
<td>72.92</td>
<td>45.78</td>
<td>40.52</td>
<td>30.12</td>
<td>1.08</td>
<td></td>
</tr>
</tbody>
</table>
Table 2

<table>
<thead>
<tr>
<th>Model</th>
<th>LnL</th>
<th>AIC</th>
<th>AICc</th>
<th>AIC weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEC</td>
<td>-119,305,4256</td>
<td>242,610,8513</td>
<td>242,974,4877</td>
<td>1,06E-08</td>
</tr>
<tr>
<td>DEC+J</td>
<td>-110,486,0209</td>
<td>226,972,0418</td>
<td>227,722,0418</td>
<td>2,63E-05</td>
</tr>
<tr>
<td>DEC*</td>
<td>-106,241,2542</td>
<td>216,482,5085</td>
<td>216,846,1448</td>
<td>0,004984487</td>
</tr>
<tr>
<td>DEC+J*</td>
<td>-103,960,0773</td>
<td>213,920,1546</td>
<td>214,670,1546</td>
<td>0,017948515</td>
</tr>
<tr>
<td>DIVA</td>
<td>-125,282,7663</td>
<td>254,565,5327</td>
<td>254,929,169</td>
<td>2,68E-11</td>
</tr>
<tr>
<td>DIVA+J</td>
<td>-123,189,5028</td>
<td>252,379,0055</td>
<td>253,129,0055</td>
<td>7,99E-11</td>
</tr>
<tr>
<td>DIVA*</td>
<td>-104,161,6852</td>
<td>212,323,3704</td>
<td>212,687,0068</td>
<td>0,039880977</td>
</tr>
<tr>
<td>DIVA+J*</td>
<td>-104,972,7468</td>
<td>215,945,4937</td>
<td>216,695,4937</td>
<td>0,006519762</td>
</tr>
<tr>
<td>BAYAREA</td>
<td>-131,223,9925</td>
<td>266,447,9849</td>
<td>266,811,6213</td>
<td>7,04E-14</td>
</tr>
<tr>
<td>BAYAREA+J</td>
<td>-113,250,3658</td>
<td>232,500,7317</td>
<td>233,250,7317</td>
<td>1,66E-06</td>
</tr>
<tr>
<td>BAYAREA*</td>
<td>-119,739,3191</td>
<td>243,478,6382</td>
<td>243,842,2745</td>
<td>6,85E-09</td>
</tr>
<tr>
<td>BAYAREA+J*</td>
<td>-100,011,714</td>
<td>206,023,4279</td>
<td>206,773,4279</td>
<td>0,930638293</td>
</tr>
</tbody>
</table>