Opinion

Tracing the origin and evolution of the orchid family through genes and trees

El origen y la evolución de la familia de las orquídeas a través de genes y árboles

With 29,524 species distributed in nearly all terrestrial ecosystems, orchids are one of the largest groups amongst the flowering plants and an important component in the floras of tropical countries (**Camara-Leret** *et al.*, 2020; **Pérez-Escobar** *et al.*, 2022). Beyond their cultural significance (**Seyler** *et al.*, 2020), orchids are far from being a trivial plant family for human societies: they represent an important biological resource in the economies of several countries, valued for their ornamental and medicinal properties (**Hinsley** *et al.*, 2017). The multiple adaptations that orchids evolved to interact with other plants, fungi, and animals (**Ackerman** *et al.*, 2023) and upon which they rely for their subsistence and reproduction have made them a group of wide scientific interest, attracting research from different fields in plant science for nearly 300 years (**Darwin**, 1862; **Dressler**, 1993; **Rudall & Bateman**, 2002).

Following the introduction of DNA sequencing in evolutionary biology studies, the evolution of orchids has been actively researched, enabling for the first time formal research on the time and place of origin of orchids, and their mode and tempo of evolution. Hundreds of studies have been conducted to elucidate relationships within the family, the majority of which have focused on specific groups at the generic and sub-tribal levels (Cameron *et al.*, 1999; Bateman *et al.*, 2003; Freudenstein *et al.*, 2004; Bateman *et al.*, 2021). These studies mostly relied on inferring phylogenetic trees derived from a handful of genes sequenced from nuclear and organellar genomic regions. They were seminal in the establishment of new taxonomic classification systems (Chase *et al.*, 2015; 2021) and the conduction of comparative phylogenetic studies on character evolution (Martins *et al.*, 2018; Bogarín *et al.*, 2019, Pramanik *et al.*, 2023; Liu *et al.*, 2024). Yet, these loci's limited informativeness and reduced taxon sampling prevented the resolution of several relationships at different taxonomic levels (Chase *et al.*, 2021).

With the advent of high-throughput sequencing, the generation of DNA sequences for hundreds of genes from both preserved and degraded specimens became scalable, facilitating the inference of orchid family-wide phylogenies (**Givnish** *et al.*, 2016; **Li** *et al.*, 2019; **Serna-Sánchez** *et al.*, 2021; **Pérez-Escobar** *et al.*, 2021; **Zhang** *et al.*, 2023), and resolving previously unsupported relationships. The first attempt to estimate the birthplace of the orchid family using molecular data was conducted by **Givnish** *et al.* (2016). This study relied on a phylogeny that sampled 201 species, which represents approximately 1% of the currently accepted species, covering all subfamilies, 18/19 tribes and 40/43 subtribes. The phylogeny was inferred from 75 plastid genes, a genomic region uniparentally inherited in orchids (**Cafasso & Cozzolino**, 2005), thus reflecting evolutionary histories that might deviate from the true species-tree relationships (**van der Niet & Linder**, 2008; **Pérez-Escobar** *et al.*, 2017; 2021). Their findings suggested that the most recent common ancestor of extant orchids inhabited the American Tropics and Australasia around 85 Ma (Figure 1 in **Givnish** *et al.*, 2016).

Limited taxon sampling can introduce biases in phylogenetic tree inferences and the estimation of ancestral characters and diversification rates (**Turner** *et al.*, 2009; **Chang** *et al.*, 2022). For this, our study (**Pérez-Escobar** *et al.*, 2024) revisited the macroevolutionary hypotheses proposed by **Givnish** *et al.* (2016) by assembling an extended phylogenomic framework of the orchid family that included 38% (338) of the currently accepted genera,

Commentary on the paper: Pérez-Escobar, O.A., et al. (2024). The origin and evolution of orchids. New Phytologist, *242*(2): 700-716 17 out of 22 subtribes, 40 out of 49 subtribes, and all subfamilies. Such a dataset was created by sequencing 353 low-copy nuclear genes from representative species of every genus; a robust phylogeny was produced (**Figure 1**) that confidently resolved most of the relationships between sampled lineages. Additionally, for the first time in the orchid tree of life, two previously unsampled genera, *Claderia* (Agrostophyliinae, Epidendroideae) and *Cooktownia* (Orchidiinae, Orchoideae), were placed.

While high-throughput sequencing has determined millions of DNA sequences for hundreds of specimens, producing and analyzing such a dataset for all currently accepted species remains cost-prohibitive and computationally time-intense. Thus, a strategy to build joint, compatible phylogenomic frameworks from different types of sequence data already available in public repositories seems pragmatic to achieve improved taxon sampling. Here, phylogenetic trees serve as the common output for either type of gene dataset, which then can be combined into a single framework. Leveraging multilocus datasets available in public sequence repositories (NCBI), we sampled sequences of the Internal Transcribed Spacer (ITS) and the maturase-K loci for 1920 species (approximately 7% of the currently accepted species). Subsequently, using molecular clock models (Bouckaert et al., 2019), we inferred absolute ages of divergence from the publicly available NCBI sequences and the 353 nuclear gene datasets. The analyses produced thousands of phylogenies with branch lengths equivalent to time units (Ma), from which we then randomly sampled ten and merged using a prune-and-graft approach. This approach allowed us to combine the information on species-level relationships provided by the sequences collected from NCBI with the genus, subtribe, and tribe-level relationships inferred by the 353 low-copy nuclear genes. Besides, by relying on multiple merged phylogenies for our comparative analyses on ancestral area and speciation (λ) rate estimation (instead of using a single consensus tree as traditionally done), we took into consideration phylogenetic uncertainty (e.g., branch length and topological variation), which can consistently affect the outcome of phylogenetic comparative methods (Cusimano & Renner, 2014).

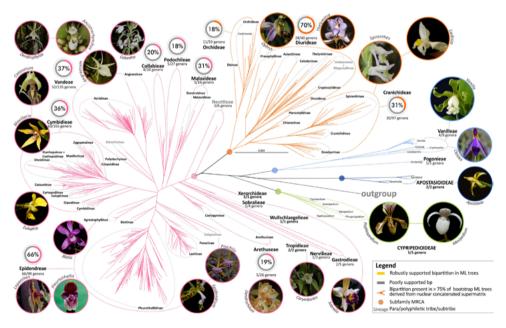


Figure 1. A consensus tree network depicting phylogenetic relationships in the Orchidaceae inferred from 200 bootstrap replicate maximum likelihood (ML) trees derived from the concatenated alignment of 339 low-copy nuclear genes. Circles at nodes represent bipartitions present in > 75% of the bootstrap ML trees. Non-monophyletic groups are highlighted in bold and grey. Samples sequenced from typological material are highlighted in bold and pink. Photos: Oscar A. Pérez-Escobar, Diego Bogarín, Sebastian Viera, Kerry Dressler. Modified from Pérez-Escobar *et al.* (2024)

Our ancestral area estimations revealed that the most recent common ancestor of extant orchids originated in Laurasia ~83 Myr ago (+/-10 Ma) (Figure 2). This result contradicts the Neotropical-Australian estimation of Givnish et al. (2016) and the hypothesis of Chase (2001) of a Gondwanan origin of the Orchidaceae, which was proposed on the family's near-cosmopolitan current distribution. During the Middle Cretaceous period, Laurasian forests were dominated by conifers and extended to high latitudes, beyond the Arctic Circle, where light was scarce for many months of the year (winter) but it was warm enough for trees to grow. The origin of orchids in Laurasia tracks that of few other monocot groups with large distribution ranges, including yams (Dioscorea) (Viruel et al., 2015) and palms (Baker & Couvreau, 2012), a family with a much richer fossil record than orchids and findings from the Northern Hemisphere dating back to the Upper Cretaceous (Estrada-Ruiz & Cevallos-Ferriz, 2009). The disparities between our results and those of Givnish et al. (2016) may stem from the influence of limited taxon sampling and phylogenetic uncertainty on ancestral character estimates, which was not accounted for in their study. Indeed, one of our ten ancestral area estimations (Figure 4 in Pérez-Escobar et al., 2024) recovered the same ancestral area obtained by Givnish et al. (2016).

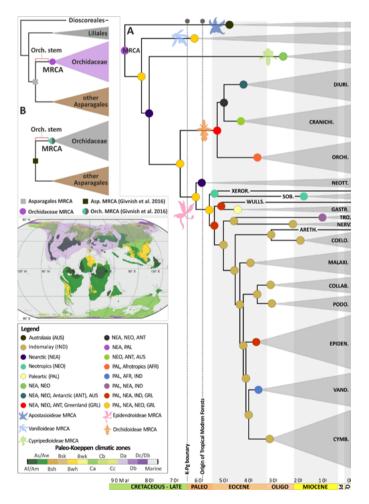


Figure 2. One of the ten possible biogeographic histories of the orchid family. A) Ancestral areas at nodes inferred on a posterior probability species tree as estimated by a Dispersal, Extinction, and Cladogenesis model (DEC). B) A summary of the outgroup sampling considered in Pérez-Escobar *et al.*, 2024 and that of Givnish *et al.*, 2016. The orchid MRCA of orchids is indicated with a circle. (Inset: A palaeoclimatic and tectonic plate reconstruction at 90–80 Ma from Burgener *et al.*, 2023, showing the possible ancestral range of the orchid MRCA as estimated by Pérez-Escobar *et al.*, 2024)

The macroevolutionary speciation dynamics of the orchid family obtained by modeling speciation rates through time in our multiple trees revealed that there was a sharp increase in speciation rates (λ) approximately 40 myrs ago (Figure 3). However, the highest number of changes in λ (from lower to higher λ) occurred mostly throughout the Miocene. The drivers behind such increases in λ rates are yet to be investigated. The estimation of tip speciation rates (i.e., predicted time at present for a lineage to undergo a speciation event) across our orchid phylogenies has, for the first time, provided a detailed map of hotspots of orchid speciation worldwide (Figure 3A). When compared with patterns of orchid species diversity, they indicated that modern speciation rates are unevenly distributed across geographical areas and might or might not be linked to species richness, a pattern known to occur in other flowering plant groups (Tietje et al., 2022). Importantly, our results pinpoint the American Tropics, especially in areas with active landscape buildup, as the area with the highest modern speciation rates on Earth. Our modern speciation rate estimations, coupled with orchid geographical distribution records, have the potential to inform conservation strategies for they can illuminate which ecosystems might have the highest evolutionary potential (Forest et al., 2007) at present. Orchids are one of the most illegally traded groups due to their high market value as an ornamental plant and one that is highly sensitive to disturbances in their native ecosystems (Parra-Sánchez & Banks-

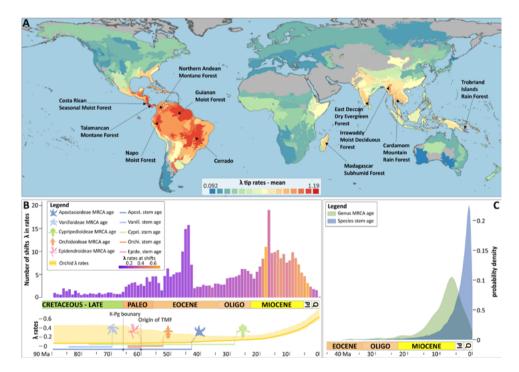


Figure 3. A) Global patterns of mean speciation (λ) tip rates per ecological region as defined by the WWF, derived from the BAMM software; warm colors indicate higher numbers of mean λ tip rates per ecological region whereas cold colors indicate lower mean λ tip rates. The highlighted geographical areas indicate eleven ecoregions with the highest mean λ tip rates (Inset: Maximum and minimum λ tip rate values for ecoregions containing 100 or more geographical distribution records for which λ tip rates were linked). Modified from **Pérez-Escobar** *et al.* (2024). **B**) Speciation dynamics of the orchid family through time. The bar plot indicates the number of shifts in speciation rates through time (in intervals of 10 Ma). The intensity of the bar color represents the average speciation rate at a given interval, with warm colors indicating higher speciation rates. The line plot at the bottom reveals variation in speciation rates through time, with the shaded area indicating confidence in the estimated rate at a given time period. **C**) The age of modern orchid diversity (species and genera) as inferred from branch lengths obtained from ten Posterior Probability species trees. Modified from **Pérez-Escobar** *et al.* (2024)

Leite, 2020). Our analyses on the age of origin of modern orchid diversity, indicated by branch lengths of the tree tips, suggest that this diversity mostly originated during the last 5 Ma (Figure 3C). This timing starkly contrasts with faster extinction rates estimated for plants (~600 species in the last 250 years) (Humphreys *et al.*, 2019). This stresses the urgent need to preserve orchid natural habitats, in particular those that might exhibit higher evolutionary potential to support greater orchid diversity.

Although assembling a comprehensive orchid tree of life remains an ongoing task that is far from complete, our study provides a solid foundation upon which a densely sampled and robust framework, built from different sequencing datasets, can be constructed in a time-efficient manner. Expanded versions of the orchid tree of life will enable the establishment of more solid and definitive taxonomic classifications and support family-wide research on the evolution of key innovations linked to the extraordinary diversification of the orchid family.

Oscar A. Pérez-Escobar¹, Diego Bogarín^{2,3}, Carlos Jaramillo⁴

¹Royal Botanic Gardens, Kew, TW9 3AE, London, UK ²Lankester Botanical Garden, University of Costa Rica, Costa Rica ³Naturalis Biodiversity Centre, Leiden, The Netherlands ⁴Smithsonian Tropical Research Institute, Panama City, Panama

References

- Ackerman, J., Phillips, R.D., Tremblay, R.L., Karremans, A., Reiter, N., Peter, C.I, Bogarín, D., Pérez-Escobar, O.A., Liu, H. (2023). Beyond the various contrivances by which orchids are pollinated: global patterns in orchid pollination biology. *Botanical Journal of the Linnean Society*, 202, 295-324.
- Baker, W.J. & Couvreur, T.L.P. (2012). Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. II. Diversification history and origin of regional assemblages. *Journal of Biogeography*, 40, 286-298.
- Bateman, R.M., Hollingsworth, P.M., Preston, J., Yi-Bo, L., Pridgeon, A.M., Chase, M.W. (2003). Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). *Botanical Journal of the Linnean Society*, 142, 1-40
- Bateman RM, Rudall PJ, Murphy ARM, Cowan RS, Devey DS, Pérez-Escobar OA. (2021). Whole plastomes are not enough: phylogenomic and morphometric exploration at multiple demographic levels of the bee orchid clade *Ophrys* sect. *Sphegodes. Journal of Experimental Botany*, 72, 654-681.
- Bogarín, D., Pérez-Escobar, O.A., Karremnans, A.P., Fernández, M., Kruzinga, J., Pupulin, F., Smets, E., Gravendeel, B. (2019). Phylogenetic comparative methods improve the selection of characters for generic delimitations in an hyperdiverse Neotropical orchid clade. *Scientific Reports*, 9, 15098.
- Bouckaert, R., Vaughan, T., Barido-Sottani, J., Duchene, S., Fourment, M., Gavryushina, A., Heled, J., Jones, G., Kuhnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Muller, N.F., Ogilvie, H.A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D.A., Siveroni, I., Suchard, M., Wu, Ch-H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PloS Computational Biology*, 15(4), e1006650.
- Burgener, L., Hyland, E., Reich, R.J., Scotese, C. (2023). Cretaceous climates: mapping paleo-Köppen climatic zones using a Bayesian statistical analysis of lithologic, paleontologic and geochemical proxies. *Palaeogeography, Palaeoclimatology, Palaeoecology, 613*, 111373.
- Caffaso, D. & Cozzolino, W.S. (2004). Chloroplast DNA inheritance in the orchid *Anacamptis* palustris using single-seed- polymerase chain reaction. *Journal of Heredity*, 96, 66-70.
- Cámara-Leret, R., Frodin, D. G., Adema, F., Anderson, C., Appelhans, M. S., Argent, G., Arias Guerrero, S., Ashton, P., Baker, W. J., Barfod, A. S., Barrington, D., Borosova, R., Bramley, G. L. C., Briggs, M., Buerki, S., Cahen, D., Callmander, M. W., Cheek, M., Chen, C. W., Conn, B J., Coode, M. J. E.... van Welzen, P. C. (2020). New Guinea has the world's richest island flora. *Nature*, 584(7822), 579-583. https://doi.org/10.1038/s41586-020-2549-5

- Cameron, K.M., Chae, M.W., Whitten, M.W., Kore, P.J., Jarrell, D.C., Albert, V.A., Yukawa, T., Hillis, H., Goldman, D.H. (1999). A phylogenetic analysis of the Orchidaceae: evidence from *rbcL* nucleotide sequences. *American Journal of Botany*, 86, 208-224.
- Chang, J., Rabosky, D.L., Alfaro, M.E. (2020). Estimating diversification rates on incompletely sampled phylogenies: theoretical concerns and practical solutions. *Systematic Biology*, 69, 602-611.
- Chase, M. (2001). The origin and biogeography of Orchidaceae. In: Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN (eds). *Genera Orchidacearum*: vol. 2. Orchidoideae (part one). Oxford University Press.
- Chase, M.W., Cameron, K.M., Freudenstein, J.V., Pridgeon, A.M., Salazar, G., van den Berg, C., Schuiteman, A. (2015). An updated classification of Orchidaceae. *Botanical Journal of* the Linnean Society, 177, 151-174.
- Chase, M.W., Schuiteman, A., Kumar, P. (2021). Expansion of the orchid genus *Eulophia* (Eulophinae; Epidendroideae) to include *Acrolophia*, *Cymbidiella*, *Eulophiella*, *Geodorum*, *Oeceoclades* and *Paralophia*. *Phytotaxa*, 491,159-168.
- Cusimano, N. & Renner, S.S. (2014). Ultrametric trees or phylograms for ancestral state reconstruction: does it matter? *Taxon*, 63(4), 721-726
- **Darwin C.** (1862). On the various contrivances by which British and foreign orchids are fertilised by insects. John Murray.
- Dressler R. (1993). Phylogeny and classification of the orchid family. Cambridge University Press.
- Estrada-Ruiz, E. & Cevallos-Ferriz, S.R.S. (2009). Palmoxylon enochii sp. nov. de la formación Olmos (Campanioano superior-Maastrichtiano inferior), Coahuila, México. Ameghiniana, 46, 577-585.
- Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Proches, S., van der Bank, M., Reeves, G., Hedderson, T.A.J., Savolainen, V. (2007). Preserving the evolutionary potential of floras in biodioversity hotspots. *Nature*, 445, 757-760.
- Freudenstein, J.V., van den Berg, C., Goldman, D.H., Kores, P.J., Molvray, M., Chase, M.W. (2004). An expanded plastid phylogeny of Orchidaceae and analysis of jackknife branch support strategy. *American Journal of Botany*, 91,149-157.
- Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J.D., Clements, M.A., Arroyo, M.T.K., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., Cameron, K.M. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proceedings of the Royal Society B: Biological Sciences, 282,* 20151553.
- Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Doucette, A., Giraldo, G., McDaniel, J., Clements, M.A., Arroyo, M.T.K., Endara, L., Kriebel, R., Williams, N.H., Cameron, K.M. (2016). Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. *Journal of Biogeography*, 43 (10),1905-1916.
- Hinsley, A., de Boer, H., Fay, M.F., Gale, S.W., Gardiner, L.M., Gunasekara, R., Kumar, P., Masters, S., Metusala, D., Roberts, D.L., Veldman, S., Wong, S., Phelps, J. (2017). A review of trade in orchids and its implications for conservation. *Botanical Journal of the Linnean Society*, 186, 435-455
- Li, Y-X., Li, Z-H., Schuiteman, A., Chase, M.W., Li, J-W., Huang, W-Ch., Hidayat, A., Wu, Sh-SH., Jin, X-H. (2019). Phylogenomics of Orchidaceae based on plastid and mitochondrial genomes. *Molecular Phylogenetics and Evolution*, 139, 106540.
- Liu, J.W., Milet-Pinheiro, P., Gerlach, G., Ayasse, M., Nunes, C.E., Alves-dos-Santos, I., Ramírez S. (2023). Macroevolution of floral scent chemistry across radiations of male euglossine bee-pollinated plants. *Evolution*, 78, 98-110.
- Martins, A., Bochorny, T., Pérez-Escobar, O.A., Chomicki, G., Monteiro, S., Smidt, E.C. (2018). From tree tops to the ground: Reversals to terrestrial habit in *Galeandra* orchids (Epidendroideae: Catasetinae). *Molecular Phylogenetics and Evolution, 127*, 952-960.
- Parra-Sánchez, E. & Banks-Leite, C. (2020). The magnitude and extend of edge effects on vascular epiphytes across the Brazilian Atlantic Forest. *Scientific Reports*, 10, 18847.
- Pramanik, D., Becker, A., Roessner, C., Rupp, O., Bogarín, D., Pérez-Escobar, O.A., Dirks-Mulder, A., Droppert, K., Kocyan, A., Smets, E., Gravendeel, B. (2023). Evolution and development of fruits of *Erycina pusilla* and other orchid species. *PLoS ONE 18*:e0286846.
- Pérez-Escobar, O.A., Gottschling, M., Chomicki, G., Condamine, F.L., Klitgaard, B.B., Pansarin, E., Gerlach, G. (2017). Andean mountain building did not preclude dispersal of lowland epiphytic orchids in the Neotropics. *Scientific Reports* 7, 4919.

- Pérez-Escobar, O.A., Dodsworth, S., Bogarín, D., Balbuena, J.A., Schley, R.J., Kikuchi, I.Z., Morris, S.K., Epitawalage, N., Cowan, R., Maurin, O., Zuntini, A., Arias, T., Serna-Sánchez, A., Gravendeel, B., Torres, M.F., Nargar, K., Chomicki, G., Chase, M.W., Leitch, I.J., Forest, F., Baker, W.J. (2021). Hundreds of nuclear and plastid loci yield novel insights into orchid relationships. *American Journal of Botany*, 108, 1166-1180.
- Pérez-Escobar, O.A., Zizka, A., Bermúdez, M.A., Meseguer, A.S., Condamine, F.L., Hoorn, C., Hooghiemstra, H., Pu, Y., Bogarín, D., Boschman, L.M., Pennington, R.T., Antonelli, A., Chomicki, G. (2022). The Andes through time: evolution and distribution of Andean floras. *Trends in Plant Science*, 27, 1-12.
- Pérez-Escobar, O.A., Bogarín, D., Przelomska, N.A.S., Ackerman, J.D., Balbuena, J.A., Bellot, S., Bühlmann, R.P., Cabrera, B., Cano, J.A., Charitonidou, M., Chomicki, G., Clements, M.A., Cribb, P., Fernández, M., Flanagan, N.S., Gravendeel, B., Hágsater, E., Halley, J.M., Hu, A-Q., Jaramillo, C., Mauad, A.V., ..., Antonelli, A. (2024). The origin and speciation of orchids. *New Phytologist*, 242, 700-716.
- Rudall, P.J. & Bateman, R.M. (2002). Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. *Biological Reviews*, 77, 403-441.
- Serna-Sánchez, M., Pérez-Escobar, O.A., Bogarín, D., Torres-Jiménez, M.F., Álvarez-Yela, A.C., Arcila-Galvis, J.E., Hall, C., de Barros, D., Pinheiro, F., Dodsworth, S., Chase, M.W., Antonelli, A., Arias, T. (2021). Plastid phylogenomics resolves ambiguous relationships within the orchid family and provides a solid timeframe for biogeography and macroevolution. *Scientific Reports*, 11, 6858.
- Seyler, B.S., Gaouse, O.G., Tang, Y., Duffy, D.C., Aba, E. (2020). Collapse of orchid populations altered traditional knowledge and cultural cultivation in Sichuan, China. *Anthropocene*, 29, 100236.
- Tietje, M., Antonelli, A., Baker, W.J., Govaerts, R., Smith, S.A., Eiserhardt, W.L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences, USA, 119*, e2120662119.
- Turner, A.H., Smith, N.D., Callery, J.A., Ebach, M. (2009). Gauging the effects of sampling failure in biogeographical analysis. *Journal of Biogeography*, 36, 612-625.
- Van der Niet, T. & Linder, P.H. (2008). Dealing with incongruence in the quest for the species tree: a case study from the orchid genus *Satyrium*. *Molecular Phylogenetics and Evolution*, 47, 154-174.
- Viruel, J., Segarra-Moragues, J.G., Raz, L., Forest, F., Wilkin, P., Sanmartín, I., Catalán, P. (2015). Late Cretaceous - Early Eocene origin of yams (*Dioscorea*, Dioscoreaceae) in the Laurasian Palaeartic and their subsequent Oligocene-Miocene diversification. *Journal of Biogeography*, 43, 750-672.
- Zhang, G., Hu, Y., Huang, M-Z., Huang, W-C., Liu, D-K., Zhang, D., Hu, H., Downing, J, L., Liu, Z-J., Ma, H. (2023). Comprehensive phylogenetic analyses of Orchidaceae using nuclear genes and evolutionary insights into epiphytism. *Journal of Integrative Plant Biology*, Vol.? pages? https://doi.org/10.1111/jipb.13462