Annals of Botany **XX**: 1–15, 2025
https://doi.org/10.1093/aob/mcaf002, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Founded 1887

ORIGINAL ARTICLE

Ancient reticulation, incomplete lineage sorting and the evolution of the pyrenoid at the dawn of hornwort diversification

Gabriel Peñaloza-Bojacá[1,*,†], J. Gordon Burleigh[2,†], Adaíses Maciel-Silva[1], D. Christine Cargill[3], David Bell[4,5], Emily B. Sessa[6], Stuart F. McDaniel[2], E. Christine Davis[2], Lorena Endara[7], N. Salazar Allen[8], Fay-Wei Li[9], Peter Schafran[9], Sahut Chantanaorrapint[10], Jeffrey G. Duckett[11], Silvia Pressel[11], Claudia Solís-Lemus[12], Karen S. Renzaglia[13,†] and Juan Carlos Villarreal A.[14,*,†]

[1]Department of Botany, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil, [2]Department of Biology, University of Florida, Gainsville, FL, USA, [3]Australian National Herbarium, Canberra, Australia, [4]Department of Botany, University of British Columbia, Canada, [5]Royal Botanic Garden, Edinburgh, UK, [6]New York Botanical Garden, New York, NY, USA, [7]Department of Biological Sciences, Clemson University, Clemson, SC, USA, [8]Smithsonian Tropical Research Institute, Panama City, Panama, [9]Boyce Thompson Institute, Cornell University, Ithaca, NY, USA, [10]Department of Biology, Faculty of Science, Prince of Songkla University, Songkhla, Thailand, [11]Natural History Museum, London, UK, [12]Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI, USA, [13]Department of Plant Biology, Southern Illinois University Carbondale, Carbondale, IL, USA, and[14]Département de Biologie, Université Laval, Quebec City, QC, Canada
[*]*For correspondence. E-mail gabriel-felipe.penaloza-bojaca.1@ulaval.ca or jcvil9@ulaval.ca*
[†]These authors contributed equally to this article.

Received: 13 July 2024 Returned for revision: 16 November 2024 Editorial decision: 8 January 2025 Accepted: 14 January 2025

• **Background and Aims** Resolving the phylogeny of hornworts is critical in understanding the evolution of key morphological characters that are unique to the group, including the pyrenoid. Extensive phylogenomic analyses have revealed unexpected complexities in the placement of *Leiosporoceros*, the previously identified sister taxon to other hornworts. We explore the role of incomplete lineage sorting (ILS) and ancient reticulation in resolving interrelationships and understanding the diversification and evolutionary processes within hornworts.
• **Methods** Using the GoFlag probe set, we sequenced 405 exons representing 234 nuclear genes, sampling 79 hornwort specimens, including representatives of all hornwort genera. We inferred the species phylogeny from gene tree analyses using concatenated and coalescence approaches, assessed ancient reticulation and ILS, and estimated the timing of divergences based on fossil calibrations.
• **Key Results** Extreme levels of gene tree incongruence challenge the sister relationship of *Leiosporoceros* to other hornworts. This phylogenetic discordance is due to ILS and ancient reticulation, the latter revealed using a network approach to identify evidence of gene flow among hornwort lineages. Hornwort diversification began in the Carboniferous with widespread family-level divergences during the mid-Cretaceous and Palaeogene.
• **Conclusions** ILS and ancient reticulation are identified as important in hornwort evolution. Patterns of hornwort diversification parallel those in other plants groups (e.g. liverworts, mosses, ferns and gymnosperms). Two scenarios on pyrenoid evolution are plausible based on the variable position of the pyrenoid-free *Leiosporoceros*. Pyrenoids were retained from a green algal ancestor and are plesiomorphic, or they evolved in response to the substantial drop in atmospheric CO_2 levels during the Carboniferous as has been hypothesized in other photosynthetic organisms. Both hypotheses require losses and gains during hornwort speciation.

Key words: Bryophytes, CO_2 levels, Cretaceous–Palaeogene boundary, evolutionary history, gene flow, hybridization, phylogenomic discordance, rapid diversification.

INTRODUCTION

With nearly 230 species worldwide, hornworts are an enigmatic lineage that is resolved as sister to the other bryophytes (Söderström *et al.*, 2016; Bechteler *et al.*, 2023). The distinctiveness of hornworts stems from a combination of unique features, such as indeterminant sporophyte growth governed by an intercalary meristem, an undifferentiated gametophyte thallus, symbiotic associations with nitrogen-fixing

cyanobacteria and the presence of a pyrenoid-based biophysical carbon-concentrating mechanism (Renzaglia *et al.*, 2009; Villarreal and Renner, 2012; Li *et al.*, 2017). Pyrenoids are found throughout eukaryotic algae (red, brown, green, chrysophytes, etc.) and have been repeatedly lost in all lineages (Atkinson *et al.*, 2016; Meyer *et al.*, 2020). The pyrenoid is a liquid–liquid phase aggregation of the enzyme Rubisco and turbocharges CO_2 to the Rubisco active site enhancing photosynthesis in aquatic environments where CO_2 diffuses more slowly than in air (Heureux *et al.*, 2017; Meyer *et al.*, 2017, 2020). Most of the knowledge of the pyrenoid assembly comes from the model green alga *Chlamydomonas reinhardtii* P.A.Dang. and includes constituent transport proteins and the protein–protein interactions that are involved in pyrenoid formation (Uniacke and Zerges, 2009; Zhan *et al.*, 2018). Hornworts are the only group of land plants possessing a pyrenoid, and the presence of this trait has intrigued biologists for centuries (Renzaglia *et al.*, 2009). The most recent reconstruction of pyrenoid evolution inferred that the trait emerged five times during hornwort diversification (Villarreal and Renner, 2012), but this analysis was based on plastid genes alone.

A recent phylogenomic study of bryophytes using 228 nuclear genes resolved many previously ambiguous nodes in liverworts and mosses (Bechteler *et al.*, 2023) but also highlighted several cases of gene tree incongruence among bryophyte lineages. The availability of a large number of genes for phylogenetic reconstruction has shed light on the complex evolutionary processes underlying the diversification of many clades. For example, conflicting gene tree topologies may result from incomplete lineage sorting (ILS) and ancient hybridization or gene flow between distinct lineages (Knowles *et al.*, 2018; Cooper *et al.*, 2023). In mosses (Guan *et al.*, 2018) and liverworts, ILS may be associated with rapid radiations, and transcriptomic data suggest a history of ancient reticulation events within liverworts (Dong *et al.*, 2022). However, the phylogenomic discordance may also arise due to either systematic error in the analyses or stochastic error in tree reconstruction (Degnan and Rosenberg, 2009). Thus, while these phylogenomic data provide new opportunities to elucidate the processes of evolution generating species diversity through time, they also present new analytical challenges that in many cases are relatively unexplored.

Molecular phylogenetic studies of hornworts have generally recovered a consistent topology that is also supported by morphology and serves as the basis for the current classification (Duff *et al.*, 2007). *Leiosporoceros* Hässel has a unique suite of morphological characters that include smooth bean-shaped spores in bilateral-alterno opposite tetrads, the presence of *Nostoc* cyanobacteria in strands that grow in parallel to the main axis of the thallus, and plastids with massive central grana but no pyrenoids (Villarreal and Renzaglia, 2006; Renzaglia *et al.*, 2007). *Leiosporoceros* has been frequently recovered as the sister to all other hornworts based on analyses of organellar sequences (Duff *et al.*, 2007; Villarreal *et al.*, 2015; Bell *et al.*, 2020) and over 400 loci for a handful of hornwort representatives (Leebens-Mack *et al.*, 2019). Based on this topology, Villarreal and Renner (2012) recovered a pyrenoid-free ancestor in all hornworts, particularly due to the fact that *Leiosporoceros* lacks pyrenoids and no other land plant outgroup possesses the trait. However, a recent analysis of 12 hornwort species and 234 nuclear loci failed to resolve *Leiosporoceros* as sister to other hornworts (Bechteler *et al.*, 2023), opening the possibility that the pyrenoid is plesiomorphic in hornworts and was secondarily lost in *Leiosporoceros*.

Here, we re-evaluate hornwort relationships using both concatenation and coalescent-based method on sequences from numerous nuclear loci across all hornwort genera. The effects of various data filtering strategies and analytical methods are assessed to infer the underlying causes of phylogenetic discordance with emphasis on the potential role of ILS and reticulation throughout the history of hornworts. Divergence times are estimated using fossil calibrations and the evolution of the hornwort pyrenoid is re-examined in the light of our phylogenomic hypotheses.

MATERIALS AND METHODS

Taxon sampling

We sampled a total of 79 hornwort specimens, including representatives of all five families and 11 hornwort genera (Söderström *et al.*, 2016). In addition, we sampled four representatives of mosses [*Sphagnum magellanicum* Brid., *Ceratodon purpureus* (Hedw.) Brid., *Pseudanomodon attenuates* (Hedw.) Ignatov and Fedosov, *Bryum argenteum* Hedw.] and twelve liverworts [*Riccardia incurvate* Lindb., *Riccardia latifrons* (Lindb.) Lindb., *Sphaerocarpos cristatus* M.Howe, *Sphaerocarpos drewei* Wigglesw., *Marchantia polymorpha* L., *Plagiochasma cordatum* Lehm. and Lindenb., *Ricciocarpos natans* (L.) Corda, *Riccia bicarinata* Lindb., *Riccia beyrichiana* Hampe ex Lehm., *Targionia hypophylla* L., *Monoclea gottschei* Lindb., *Monoclea forsteri* Hook.] as outgroup taxa.

DNA extraction

We extracted DNA using an E.Z.N.A. DNA (Omega, Bio-Tek, USA) extraction kit and Gel Extraction Kit (QIAquick) according to the manufacturer's protocol. In addition, the modified cetyltrimethylammonium bromide (CTAB) extraction protocol (Doyle and Doyle, 1987) described by Breinholt *et al.* (2021) and Bechteler *et al.* (2023) was used for most of the samples. This protocol includes lysing the cells by centrifuging them and washing with two rounds of 24:1 (v/v) chloroform–isoamyl alcohol, followed by cold isopropanol precipitation and a 70 % (v/v) ethanol wash. Next was added 2 µL of 10 mg mL^{-1} RNase A (Qiagen, Valencia, CA, USA) to remove RNA contamination between chloroform washes.

Target enrichment and sequencing assembly

We assembled a combination of published target enrichment data from Breinholt *et al.* (2021) and Bechteler *et al.* (2023) and data that are new to this study. Data from the 16 samples (14 hornworts, one moss and one liverwort; Supplementary Data Table S1) first published in Breinholt *et al.* (2021) were generated using the GoFlag 451 flagellate land plant probe set, which covers 451 exons found in 248 single- or low-copy nuclear genes. These unlinked genes appear to be scattered across hornwort chromosomes from species in all clades, based on the genomes *Anthoceros agrestis*

Paton, *Leiosporoceros dussii* (Steph.) Hässel, *Phaeoceros carolinianus* (Michx.) Prosk. and *Phaeomegaceros chiloensis* (Steph.) J.C.Villarreal (P. Schafran, pers. comm.). The data from the nine samples (five hornworts, two mosses and two liverworts; Table S1) first published in Bechteler *et al.* (2023), as well as the 70 samples (61 hornworts, three mosses and six liverworts; Table S1) new to this study were generated using the GoFlag 408 flagellate land plant probe. The GoFlag 408 probe set is a subset of the GoFlag 451 probe set. Library construction, target enrichment and sequencing were done by RAPiD Genomics (Gainesville, FL, USA) using protocols described in Breinholt *et al.* (2021). The enriched, pooled libraries were sequenced on an Illumina HiSeq 3000 platform (Illumina; 2 × 100 bp), and the paired-end raw reads are available in the NCBI SRA database (Table S1).

We assembled sequence alignments for the target regions covered by the probe sets (i.e. conserved exons) from the raw sequence data using the GoFlag pipeline described by Breinholt *et al.* (2021; see script 1, this and all scripts are in the online repository). In cases where there appears to be greater than simple allelic variation at a locus, the pipeline may retain more than one sequence from a sample. In these cases, to minimize the inclusion of possible paralogues, we removed all copies from these samples in the locus alignment. Next, we pruned columns in the alignments that had fewer than ten nucleotides from the output PHYLIP files. We also concatenated the alignments from any exons (i.e. loci) that were found in the same gene (see 'GFvs1kp.txt' in script 1, which maps GoFlag loci to their 'single-copy' 1KP gene). As a result, we obtained 234 alignments, each representing a single gene and containing between one and nine exons. These alignments were used to build 'gene' trees. Combining loci from the same gene implicitly assumes that there is no recombination within the gene; however, longer gene alignment may also reduce stochastic errors in gene-tree inference (Breinholt *et al.*, 2021; Bechteler *et al.*, 2023).

Gene tree and species tree estimation

We inferred maximum likelihood (ML) gene trees from each of 234 genes using RAxML v.8.2.9 (Stamatakis, 2014) with the GTR+Γ model with 100 random starting points. Statistical confidence of each gene tree was assessed by performing 100 bootstrap (BP) replicates (script 2). To estimate the species tree, we ran ASTRAL-III, which uses a coalescent-based approach to estimate the species tree (Zhang *et al.*, 2018), using all 234 ML gene trees from the RAxML analyses as input. We assessed support for the ASTRAL topology based on the local posterior probabilities (LPP; Sayyari and Mirarab, 2016) for the main topology (q1), the first alternative topology (q2) and the second alternative topology (q3; script 3).

We also inferred an ML phylogenetic tree from a concatenated supermatrix of all gene alignments using IQ-TREE v.2.1.3 (Nguyen *et al.*, 2015), with '-S option'. We partitioned the supermatrix alignment by gene and used ModelFinder (Kalyaanamoorthy *et al.*, 2017) to automatically select the best substitution model for each partition. Branch support was assessed using ultrafast bootstrapping (1000 replicates; script 4; Hoang *et al.*, 2017).

Measuring concatenation and coalescent-based topological signal

Since our analyses showed discordant topologies in the branch leading to *Leiosporoceros dussii*, we applied a method to assess and compare inconsistent genes between concatenation-based IQ-TREE (T1) and Quartet-based ASTRAL (T2; Shen *et al.*, 2021). For incongruent internal bipartition(s) between T1 and T2, we defined a concatenation-based gene-wise phylogenetic signal as the difference in gene-wise log-likelihood score (ΔGLS) for T1 versus T2 and a quartet-based gene-wise phylogenetic signal as the difference in gene-wise quartet score (ΔGQS) for T1 versus T2 (see Supplementary Data; Shen *et al.*, 2021). ΔGLS and ΔGQS values can be positive, negative or zero. We assessed the consistency of gene-wise phylogenetic signal by calculating the following two measures for every gene: -**consistent**: genes whose ΔGLS > 0 and ΔGQS > 0 (from T1) and genes whose ΔGLS < 0 and ΔGQS < 0 (from T2); and -**inconsistent**: genes whose ΔGLS ≥ 0 and ΔGQS ≤ 0 or vice versa (Shen *et al.*, 2021). We also examined the topologies using two different filtered datasets. First, we filtered out genes that did not include a sequence from *Leiosporoceros*, resulting in a dataset we called Only Leios, which included 195 genes. Additionally, we created separate datasets to include only consistent genes and only inconsistent genes (Supplementary Data Table S2; script 5).

Branch support and concordance analyses

Branch support was measured using bootstrap and LPP in IQ-TREE and ASTRAL, respectively. Additionally, we performed concordance analyses in the following datasets: (1) main dataset (Full data with 234 genes and Only Leios with 195 genes); (2) subset of data (Full data consistent with 165 genes and Only Leios consistent with 133 genes); and (3) divergent topologies (concatenation-based IQ-TREE, T1 and Quartet-based ASTRAL, T2). We calculated gene (gCF) and site concordance factors (sCF) to investigate topological conflict around each branch of the species tree in IQ-TREE with options '–gcf and –scf ' (Minh *et al.*, 2020). In addition, IQ-TREE estimates gDFP, a gene discordant factor due to paraphyly or the gene discordance factor due to lack of information in the genes of the quartet (script 6).

These analyses were conducted on every branch of the species tree. gCF and sCF represent the percentage of decisive gene trees and sites. gCFs/sCFs were categorized as follows: weak < 33 %, moderate 33 %–50 %, strong > 50 % (following Minh *et al.*, 2020; Bechteler *et al.*, 2023; Cooper *et al.*, 2023). We performed exploratory analyses to assess whether gene tree patterns are consistent with the neutral ILS model using IQ-TREE. To do this, a χ^2-test was used to determine whether the frequency of gene trees (gCF) and sites (sCF) supporting the alternative topologies was significantly different (Chan *et al.*, 2020). Under the assumption of ILS, the discordant topologies should be supported by an approximately equal number of gene trees or sites, which would result in a non-significant χ^2 (Minh *et al.*, 2020). Thus, rejecting the χ^2 suggests that processes other than ILS, including gene tree error, may be contributing to the discordance. For this test we followed the approach of Chan *et al.*(2020) such that the χ^2-test was performed in R

(Supplementary Data Table S3; Lanfear *et al.*, 2018; Chan *et al.*, 2020; R Core Team 2024).

Alternative topologies, especially the position of *Leiosporoceros*, were tested using tree topology tests (with -wpl to print partition-wise log likelihoods for both trees) and testing the constrained tree in IQ-TREE (script 7). We used several tree topologies tests in IQ-TREE using the RELL approximation (Kishino *et al.*, 1990) including bootstrap proportion (BP), Kishino–Hasegawa test (Kishino and Hasegawa, 1989), Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 1999), expected likelihood weights (Strimmer and Rambaut, 2002) and approximately unbiased (AU) test (Shimodaira, 2002). We then identified the most influential genes using partition-wise log likelihoods (-wpl) to identify which genes were contributing most phylogenetic signal towards one tree but not the other. IQ-TREE looks at the gene-wise log-likelihood (logL) differences between the two given trees T1 and T2. Genes that have the largest logL(T1) − logL(T2) will favour T1, whereas genes showing the largest logL(T2) − logL(T1) will favour T2 (Supplementary Data Table S4; Minh *et al.*, 2018, 2019).

Network inferences

We used the Species Networks applying the Quartets (SNaQ) method within PhyloNetworks (Solís-Lemus and Ané, 2016; Solís-Lemus *et al.*, 2017) to examine the contribution of ILS and reticulation to the phylogenetic history of hornworts. This package uses maximum pseudolikelihood to fit a network while also accounting for ILS. SNaQ uses concordance factors (CFs), which are the frequencies of the three possible unrooted topologies of each set of four taxa (i.e. quartets) in a sample of gene trees (Solís-Lemus *et al.*, 2017), as input. Because these analyses are computationally demanding on large datasets, we randomly subsampled at the generic level (55 species from all genera) of hornworts. We then ran SNaQ allowing the maximum number of reticulations to vary from $h = 0$ to $h = 4$ and selected the best h using a slope heuristic suggested by Solís-Lemus and Ané (2016). As input, we used the 234 gene trees, and the T1 (concatenation-based IQ-TREE) and T2 (coalescence-based ASTRAL) to limit noise introduced by missing data. Each SNaQ run was performed with ten optimization iterations (script 8). In addition, we used Hybrid Detector (HyDe) to detect the phylogenetic invariants using a coalescent model with hybridization to infer probability of hybridization of three ingroup taxa relative to an outgroup taxon in hornworts (Blischak *et al.*, 2018). The parameter γ represents the probability that gene trees with a hybrid population sister to parent X would arise under the parental population trees, whereas $1 - \gamma$ is the probability of a hybrid population being sister to parent Y (script 9).

Divergence time estimation

Divergence time estimation was performed using treePL (Smith and O'Meara, 2012). The ML tree was the input for treePL based on the Full data dataset (234 genes) and gene trees previously generated with RAxML (see above) for concatenation-based IQ-TREE (T1) and coalescence-based ASTRAL (T2). All trees were rooted in the outgroup taxa using the program pxrr in phyx (Brown *et al.*, 2017) before being used in the dating analyses (script 10). Priming and cross-validation analysis was performed using the best ML tree and all four calibrations. Best optimization parameters for concatenation-based analyses and quartet-based analyses were as follows: for IQ-TREE (T1), 'opt=2, moredetail; optad=2, moredetailad; optcvad=1, moredetailcvad'; and for ASTRAL (T2), 'opt=5; optad=5; optcvad=3, moredetailcvad'. Cross-validation was conducted five times using these parameters and indicated stable values of a smoothing parameter = 10. To obtain confidence intervals on the dated tree, the treePL analysis was run with the bootstrap replicates using the same calibration, optimization and cross-validation values as outlined above. Trees were visualized in FigTree v.1.4.3 (Rambaut, 2017).

The fossil record in hornworts is sparse, and earliest reliable fossils are only from the Upper Cretaceous (Villarreal and Renner, 2012). We constrained the root to 515 (±10) Ma, using a normal distribution based on previous studies (Feldberg *et al.*, 2021; Ignatov and Maslova, 2021; Bechteler *et al.*, 2023). The age of *Anthoceros* L. spore type A from the Baqueró Formation, Argentina (118.56 ± 3.7 Ma; Archangelsky and Villar de Seone, 1996), was used to constrain the stem node of *Anthoceros* and the rest of the hornworts. The spore exine is ornamented with warts on the proximal face resembling two extant Neotropical species, *Anthoceros tuberculatus* Lehm. et Lindenb. and *Anthoceros tristanianus* J.C.Villarreal, J.J.Engel et Váňa (Peñaloza-Bojacá *et al.*, 2020). We used the age of *Notothylites nirulai* from the Deccan Intertrappean beds of Mohgaonka, India (Maastrichtian, 65–70 Ma; Chitaley and Yawale, 1980), to constrain the age of the stem node of *Notothylas* Sull. ex A. Gray and *Phaeoceros* Prosk. (including *Paraphymatoceros* Hässel). The petrified fossil of an entire plant has similar thallus size, sporophyte size, lack of stomata in sporophyte and elater shape as extant *Notothylas*. The third calibration was based on a fossil assigned to *Phaeoceros* sp. from the Uscari Formation, Costa Rica (Lower Miocene, 15–23 Ma; Graham, 1987). This spore fossil has six depressions in its distal face and closely resembles extant *Phaeomegaceros fimbriatus* (Gottsche) R.J.Duff, J.C.Villarreal, Cargill et Renzaglia in size and ornamentation (Villarreal and Renner, 2012). For the liverwort outgroup, we incorporated the fossil *Ricciopsis ferganica* dating to 227 Ma (Moisan *et al.*, 2012). Additionally, we used two constraints, 448 Ma for liverworts and 417 Ma for mosses, following the results of Bechteler *et al.* (2023). We report mean ages and the 95 % highest posterior density (HPD) values.

RESULTS

Phylogenomic inferences

The ML concatenated supermatrix and the coalescence-based analyses from ASTRAL-III using sequences from 234 genes, containing 80 855 total sites, yielded mostly congruent topologies. Four of the five hornwort families had similar topologies between analyses with BS support = 100 % (Supplementary Data Fig. S1) and LPP= > 0.9 (Fig. S2). The families recovered are: Dendrocerotaceae J.Haseg. [including *Dendroceros* Nees, *Megaceros* Campb., *Nothoceros* (R.M.Schust.) J.Haseg. and *Phaeomegaceros* R.J.Duff]; Phymatocerotaceae R.J.Duff

(including only *Phymatoceros* Stotler); Notothyladaceae Müll. Frib. ex Prosk. (including *Notothylas*, *Paraphymatoceros* and *Phaeoceros*); and Anthocerotaceae Dumort (including *Anthoceros* and *Folioceros* D.C.Bharadwaj). However, the position of the monotypic genus *Leiosporoceros* differs in the two analyses. ML analysis places *Leiosporoceros* as sister to the rest of the hornworts (BS support = 100 %; Fig. 1, T1). ASTRAL-III analyses place *Leiosporoceros* as sister to the family Anthocerotaceae (LPP = 0.7; Fig. 1, T2). Another alternative topology was observed with the coalescence and concatenation analyses (T3), with the family Anthocerotaceae identified as the sister group to the clade consisting of *Leiosporoceros* and the other hornwort families (Fig. 1; T3). This topology exhibited lower support values (LPP < 0.3; BS < 70 %) at the node corresponding to *Leiosporoceros*, and tree topology tests did also not yield significant values supporting this topology (see below). Therefore, our analyses were centred on making comparisons between T1 and T2.

Gene inconsistency in concatenation and coalescent approaches

To assess the origin of the inconsistency in the placement of *Leiosporoceros*, we quantified the distribution of concatenation-based phylogenetic signal (T1: *Leiosporoceros* sister to all hornworts) with quartet-based phylogenetic signal (T2: *Leiosporoceros* sister to Anthocerotaceae) for every dataset (see Table 1). The phylogenomic matrices, along with the ΔGLS and ΔGQS values for each gene, are provided in Supplementary Data Table S2. We found that when analysing our full data (234 genes) using concatenation-based methods, the proportion of genes recovering topology T1 (Fig. S1) was higher (55.9 %) than that recovering topology T2 (44.1 %). However, in the quartet-based analysis, more genes recovered T2 (52.6 % of the 234 genes) than T1 (47.4 %). Similar results were obtained when excluding genes in which *Leiosporoceros* was not present (Only Leios: 195 genes; Table S2). The concatenation approach supported T1 with 107 genes (54.5 %; Fig. S3), while the coalescence approach favoured T2 with 98 genes (51.8 %; Fig. S4).

The phylogenetic gene signals between concatenation-based (T1) and quartet-based (T2) approaches showed similar results. For the full dataset, 161/234 (68.8 %) of the genes were consistent; that is, their ΔGLS and ΔGQS values had the same signs. On the other hand, we found that 73/234 (31.2 %) of the genes were not consistent (ΔGLS and ΔGQS values had opposite signs). For the Only Leios dataset, 133/195 (68.2 %) of the genes were consistent, and 62/195 (31.8 %) of the genes were not consistent.

Phylogenetic discordance

T1 (*Leiosporoceros* sister to all hornworts) was inferred with the full dataset only with concatenation-based method in IQ-TREE and presented high values of support (BS = 100 %) and congruence (gCF > 90 % and sCF > 50 %; Supplementary Data Fig. S5). IQ-TREE analyses with only the consistent genes produced the same result (Fig. S6). In contrast, datasets with T2 (*Leiosporoceros* sister to Anthocerotaceae) as the main topology, inferred with quartet-based method in ASTRAL (Fig. S7) and base species tree inferred with coalescence-based method (ASTRAL-III) vs Base tree genes inferred with the concatenation method (IQ-TREE; Fig. S8), presented low levels of support (LPP < 0.7) and high levels of discordance (gCF and sCF < 37 %; Figs 1 and S7), even with consistent genes (Figs S9 and S10). Similar results were obtained from ASTRAL analyses with the Only Leios genes (Figs S4 and S11) and their consistent genes (Figs S12 and S13).

Despite the difference between T1 and T2, and the discrepancy values between the datasets, these were not significantly different ($\chi^2 > 0.05$) in the nodes leading to *Leiosporoceros* (Supplementary Data Table S3). For T1, gDF1 and gDF2 were not significantly different ($\chi^2 > 0.05$), and we observed a similar result to sDF1 and sDF2. This is consistent with high levels of ILS in the gene trees. Furthermore, when exploring the *Leiosporoceros* node for T2 across the different datasets, we observed that in most cases, gDF1 and gDF2 were not significantly different ($\chi^2 > 0.05$). However, in the coalescence and consistent gene analyses, gDF1 and gDF2 were significantly different, so we cannot reject the null hypothesis of equal frequencies of the alternate topologies ($\chi^2 > 0.05$), suggesting evidence of reticulation. These observations for sDF1 and sDF2 were not significantly different in all datasets for T2. This confirms the presence of ILS and reticulation in the gene trees.

The three topology tests indicated a slight preference for T1 while rejecting T2 and T3 (Table 2). Notably, the most influential genes yielded similar results for T1 and T2 topologies. In our full dataset of 234 genes, we observed T1 = 90 genes, T2 = 78 genes and T3 = 62 genes. For the subset of consistent genes (161 in total), the counts were T1 = 61 genes, T2 = 63 genes and T3 = 36. In the Only Leios dataset of 195 genes, we found T1 = 79 genes, T2 = 63 genes and T3 = 52. Furthermore, among the consistent genes in the Only Leios group (133 genes), we recorded T1 = 51 genes, T2 = 55 genes and T3 = 27. These results suggest that the set of genes used may have a similar phylogenetic signal for both T1 and T2. Another node that displayed a degree of incongruence was the one leading to *Notothylas*, *Paraphymatoceros* and *Phaeoceros*. This node resolved the same topology using both concatenation and coalescent-based methods although the gDFP values were close to 30 %, indicating potential rates of paraphyly or insufficient information in the genes (nodes G, H and I in Fig. 1, respectively; Supplementary Data Figs S3 and S5–S13).

Reticulation events in hornworts

We explored phylogenetic networks with different numbers of reticulation (hybridization) events (h = 0–3; Supplementary Data Figs S14 and S15). The best scores were two reticulations events for T1 (-Ploglik = 255.42; Fig. S16) and three for T2 (-Ploglik = 254.74; Fig. S17), suggesting reticulation events in the evolutionary history of hornworts (Fig. 2).

Two hybridization events were recovered based on T1 (Fig. 2A; Supplementary Data Table S5). The first was in *Nothoceros* with $\gamma = 0.115$ (gamma value), showing a low hybridization signal, and the stem branch leading to *Leiosporoceros*, *Anthoceros* and *Folioceros* with $\gamma = 0.885$, showing a high hybridization signal. The second event occurred between *Leiosporoceros* with $\gamma = 0.0292$ and *Folioceros* with $\gamma = 0.971$,

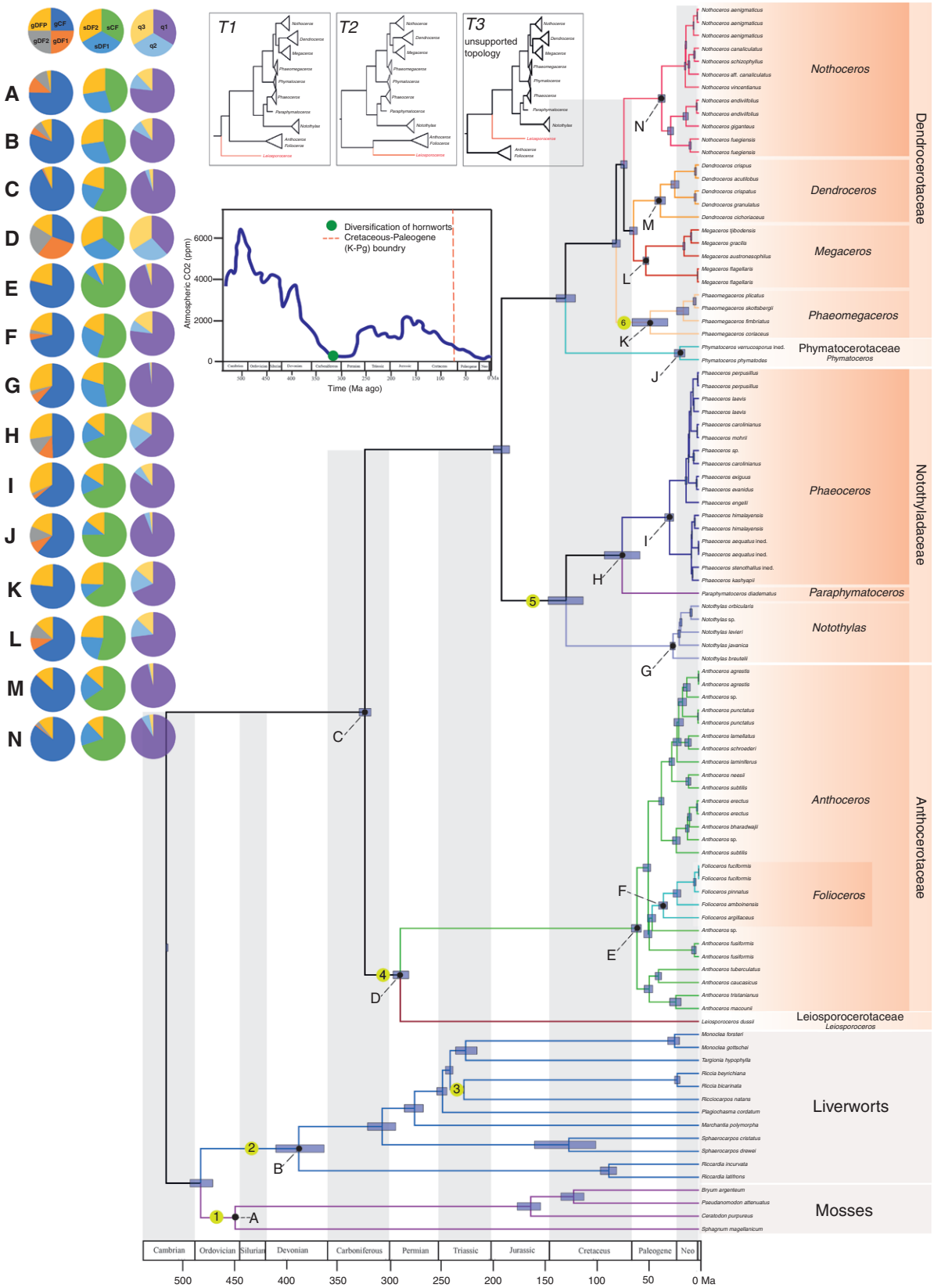

FIG. 1. Divergence time estimates of hornwort genera based on an ASTRAL analysis of 234 nuclear genes. The nodes of interest (highlighted by major clade colour) have been labelled (A–N) to illustrate quartet values from gene concordance factor (gCF) values (left) and site concordance factor (sCF) values (centre). gCF and sCF represent the percentage of decisive genes and sites at each branch, respectively. Pies for quartet values are from concordance factor values: CF (topology shown) and alternative options (DF1, DF2, DFP) and site concordance factor values (sCF, sDF1, sDF2); see Fig. S9 and Table S4. Quartet values from

suggesting a possible hybridization event in *Folioceros*, having *Leiosporoceros* and *Anthoceros* as parental groups. On the other hand, when observing the Quartet CF (observed vs. expected) derived from the gene trees of the SNaQ hybridization analysis, we identified potential discrepancies in the phylogenetic relationships of quartets within the gene trees compared to the phylogenetic network, particularly concerning taxon interactions involving *Leiosporoceros* (Fig. 2C). The results suggest discrepancies or possible hybridization events involving this genus in the gene trees compared to the phylogenetic network.

Likewise, three reticulation events were detected in the analyses based on T2 (Fig. 2B; Supplementary Data Table S5). The first was among *Notothylas* with a value of $\gamma = 0.069$ and *Phaeoceros* with a value of $\gamma = 0.931$, suggesting a possible hybridization event with *Notothylas* and *Paraphymatoceros* as parental groups. The second event occurred between the *Dendroceros* and *Megaceros* clade with a value of $\gamma = 0.0929$ and the *Leiosporoceros*, *Anthoceros* and *Folioceros* clade with a value of $\gamma = 0.907$. The third event was recorded in *Anthoceros* with a value of $\gamma = 0.972$ and a possible extinct or unsampled taxon with a value of $\gamma = 0.0278$. *Anthoceros* and this species possibly presented a case of reticulated evolution with *Leiosporoceros* and an extinct taxon. Additionally, Quartet CF (observed vs expected) showed reasonable agreement between the gene trees and the phylogenetic network (Fig. 2C), suggesting a strong correlation between these two variables. Despite limited conflict in network analyses and quartet scoring, gene flow analysis using HyDe showed extensive gene flow among hornwort genera, and there were only a few pairs of taxa between which no gene flow was identified (Table S6).

Table 1. *Main hornwort topologies. T1:* Leiosporoceros dussii *sister to the other families of hornworts (ML concatenation method); T2:* Leiosporoceros dussii *is the sister group of Anthocerotaceae (ASTRAL coalescent method); T3: Anthocerotaceae sister to the clade of* Leiosporoceros dussii *and other families of hornworts (alternative topology in coalescence and concatenation methods). The T3 topology does not have support at the* L. dussii *node. The full dataset includes the 234 genes. In contrast,* Leiosporoceros *lacks 39 genes (Villarreal et al., 2018) and we created the Only Leios dataset and performed similar analyses.*

Data set	Number of genes	Genes supporting each topology		
		T1	T2	T3
Full data	234 (100 %)	90 (38.5 %)	78 (33.4 %)	62 (26.5 %)
Full data consistent genes	161 (100 %)	61 (37.8 %)	63 (39.1 %)	36 (22.4 %)
Only Leios	195 (100 %)	79 (40.5 %)	63 (32.3 %)	52 (26.7 %)
Only Leios consistent genes	133 (100 %)	51 (38.4 %)	55 (41.3 %)	27 (20.3 %)

Table 2. *Tree topology tests. T1:* Leiosporoceros dussii *is sister to the other families of hornworts; T2:* Leiosporoceros dussii *is the sister group of Anthocerotaceae; T3: Anthocerotaceae sister to the clade of* Leiosporoceros dussii *and other families of hornworts; DeltaL: logL difference from the maximal logL in the set; bp-RELL: bootstrap proportion using RELL method; p-KH:* P*-value of one-sided Kishino–Hasegawa test; p-SH:* P*-value of Shimodaira–Hasegawa test; c-ELW: expected likelihood weight; p-AU:* P*-value of approximately unbiased (AU) test. Plus signs denote the 95 % confidence sets and minus signs denote significant exclusion. All tests were performed with 10 000 resampling using the RELL method in IQ-Tree.*

	Data set	Topology	logL	deltaL	bp-RELL		p-KH		p-SH		c-ELW		p-AUU	
Full data		T1	−1202258.58	0	1	+	1	+	1	+	1	+	1	+
Full data		T2	−1202466.911	208.33	0	−	0.0001	−	0.0001	−	1.68e-07	−	1.46e-05	−
Full data		T3	−1202511.488	252.91	0	−	0	−	0	−	1.19e-37	−	4.81e-64	−
Full data	Consistent genes	T1	−848377.7561	0	0.998	+	0.997	+	1	+	0.998	+	0.997	+
Full data	Consistent genes	T2	−848511.2003	133.44	0.002	−	0.0027	−	0.0031	−	0.00213	−	0.00308	−
Full data	Consistent genes	T3	−848604.8049	227.05	0	−	0	−	0	−	1.16e-25	−	5.78e-48	−
Only Leios		T1	−1015587.001	0	0.953	+	0.947	+	1	+	0.952	+	0.96	+
Only Leios		T2	−1015632.14	45.14	0.047	−	0.053	+	0.063	+	0.0479	−	0.0429	−
Only Leios		T3	−1015673.389	86.39	0	−	0	−	0	−	3.43e-09	−	6.79e-13	−
Only Leios	Consistent genes	T1	−743956.0191	0	0.999	+	0.999	+	1	+	0.999	+	0.999	+
Only Leios	Consistent genes	T2	−744114.4576	158.44	0.0006	−	0.0006	−	0.0006	−	0.000635	−	0.000994	−
Only Leios	Consistent genes	T3	−744189.0778	233.06	0	−	0	−	0	−	5.51e-33	−	0.000309	−

ASTRAL analyses are presented as pie-charts (right). ASTRAL pies are divided into q1 or topology shown (purple), q2 (blue, 2nd alternative hypothesis) and q3 (orange, 3rd alternative hypothesis) with the percentage for q1 included in the pie diagram; see Appendix 5, Figs S20 and S21. The detailed chronogram with node heights represents mean ages and bars the 95 % highest posterior density intervals reported in Appendix 5, Fig. S19. Numbers represent the calibrations: 1: divergence of mosses, 2: divergence of liverworts, 3: *Ricciopsis ferganica* fossil, 4: *Anthoceros* spore type A; 5: *Notothylites nirulai* fossil; and 6: fossil assigned to *Phaeomegaceros* sp. (inset). Average values of atmospheric CO_2 (ppm) levels during the last 500 Ma (Berner, 2001; Badger *et al.*, 2002; Renne *et al.*, 2013; Steinthorsdottir and Vajda, 2015).

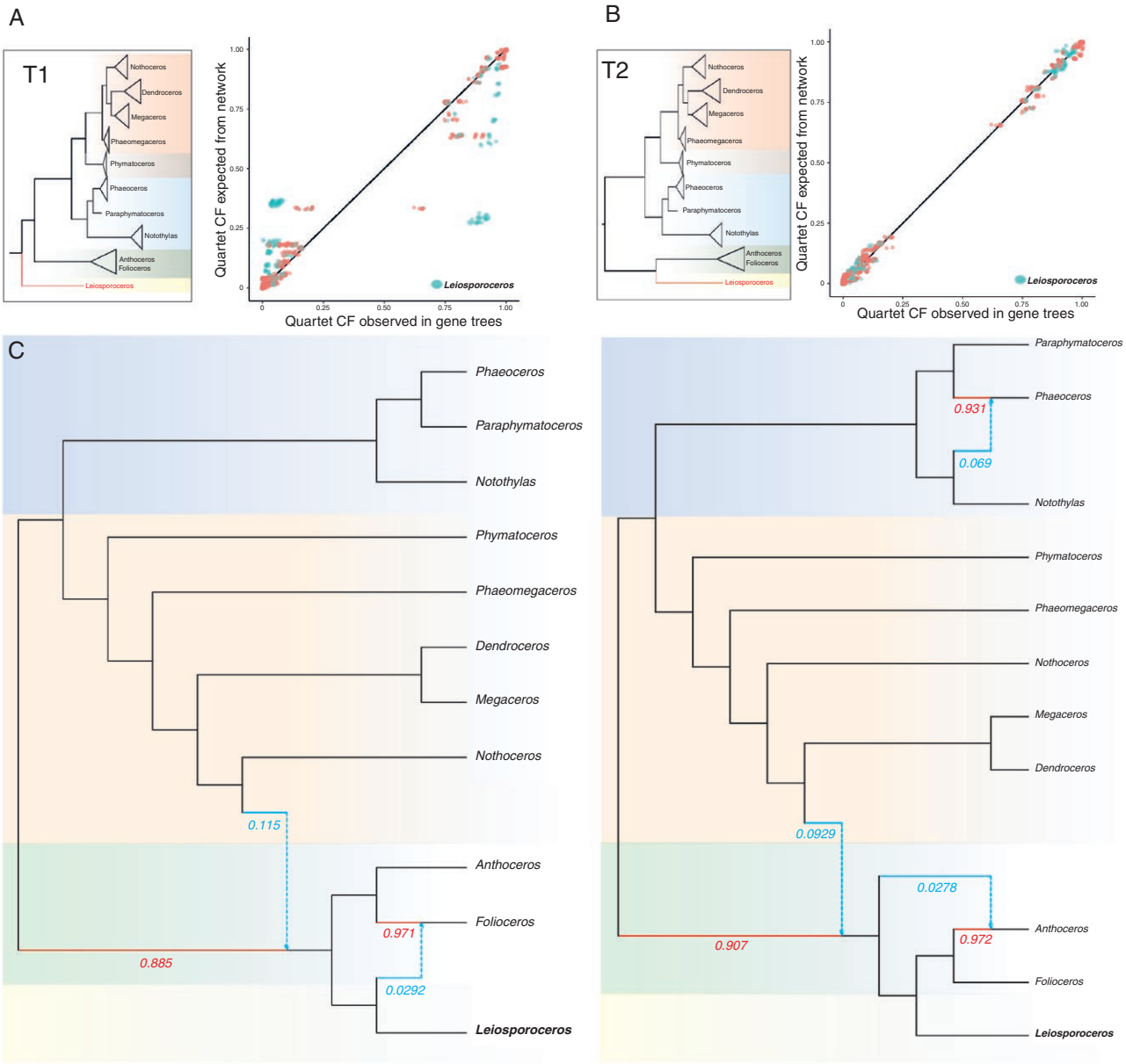

FIG. 2. Estimates of ancient reticulate evolution (SNaQ) from all hornwort genera using 234 genes and number of hybridizations (h_{max} = 3). (A) Estimated network based on topology T1 (*Leiosporoceros* sister to all hornworts) with best scores across hybrid values, –Ploglik = 255.42. (B) Estimated network based on topology T2 (*Leiosporoceros* sister to Anthocerotaceae) with best scores across hybrid values, –Ploglik = 254.74. On the left corner, we present the recovered topology and to the right the correlation graphic with the observed concordance factors versus the expected concordance factors. Light blue quartets include *Leiosporoceros*. (C) Best estimated network with possible reticulation events based on topologies T1 and T2. Red and blue lines indicate hybrid edges. Red and blue numbers indicate estimated inheritance probabilities from major and minor parental species, respectively. See Supplementary Data Figs S14 and S15 and Table S5.

Hornworts divergence times

Two topologies (T1 and T2) were dated to explore the alternative evolutionary timelines of hornworts implied by the different trees. The crown age of hornworts ranges from 337.6 Ma (HPD 330.7–348.5 Ma) for T1 to 322.5 Ma (HPD 316.8–328.2 Ma) for T2 (Supplementary Data Figs S18 and S19 for HPD values for all clades) in the Carboniferous. Both chronograms showed similar divergence ages for most taxonomic groups (Fig. 1). For the orders we estimated a stem age in the Carboniferous and Permian for Leiosporocerotales Hässel T1: 337.6 Ma (HPD: 330.7–348.6 Ma) and T2: 288.3 Ma (HPD: 316.8–328.2 Ma); Anthocerotales Limpr. T1: 299.4 Ma (HPD: 291.9–307.5 Ma) and T2: 288.3 Ma (HPD: 280.3–295.5 Ma) originated in the Permian. Notothyladales Hyvönen et Piippo and Dendrocerotales Hässel diversified in the Jurassic T1: 182.1 Ma (HPD: 170.7–193.5 Ma) and T2: 190.5 Ma (HPD: 182.9–198.6 Ma), and Phymatocerotales R.J.Duff T1: 129.3 Ma (HPD: 120.1–139.5 Ma) and T2: 129.2 Ma (HPD: 119.4–138,0 Ma) in the Cretaceous. Finally, most genera diversified (crown age) in the Palaeogene (Fig. 1; Figs S18 and S19), from HPD 24–58 Ma in T1 to HPD 25–59 Ma in

T2. *Paraphymatoceros* has an origin between the Cretaceous and Palaeogene T1: 73.6 Ma (HPD: 57.4–93.2 Ma) and T2: 73.9 Ma (HPD: 57.0–91.5 Ma). *Phymatoceros* has a Neogene origin T1: 17.8 Ma (HPD:14.3–23.2 Ma) and T2: 17.9 Ma (HPD: 13.5–23.8 Ma).

DISCUSSION

Prior to the advent of molecular phylogenetics, the systematics and classification of hornworts were widely debated and subject to conflicting interpretations (Hasegawa, 1988; Hässel de Menéndez, 1988). Following the first comprehensive molecular analysis across hornwort diversity (Duff *et al.*, 2007), a consensus emerged that challenged existing concepts of hornwort relationships. The widely recognized five or six genera were segregated into ten (11) genetically distinct taxa. *Leiosporoceros* surfaced as the sister taxon to the other hornworts, and the distinct morphological features of the genus provided support for this position (Villarreal and Renzaglia, 2006; Renzaglia *et al.*, 2007). Our analysis of sequences from 234 nuclear genes from 60 hornwort species, which includes representatives of all genera, is consistent with previous studies using organellar loci with one important exception (Duff *et al.*, 2007; Villarreal *et al.*, 2015), namely incongruence with respect to the placement of the branch leading to *Leiosporoceros*. We have shown that the position of *Leiosporoceros*, at the dawn of hornwort diversification, is influenced by ILS and ancient reticulation events.

Gene tree/species tree conflict and ILS in hornworts

Most previous studies based on a few organellar loci with high levels of hornwort sampling (Duff *et al.*, 2007; Renzaglia *et al.*, 2009; Villarreal and Renner, 2013; Villarreal and Renzaglia, 2015; Villarreal *et al.*, 2015) or low levels of hornwort sampling with a large number of markers (Leebens-Mack *et al.*, 2019; Breinholt *et al.*, 2021; Shen *et al.*, 2024) recovered *Leiosporoceros* as the sister to other hornworts. The quite distinct gametophyte and sporophyte morphology of *Leiosporoceros* support this phylogenetic placement. However, Duff *et al.* (2007, their fig. 11) removed 193 edited sites from *rbc*L and *nad*5 and recovered *Leiosporoceros* as sister to Anthocerotaceae, suggesting that the widely accepted position

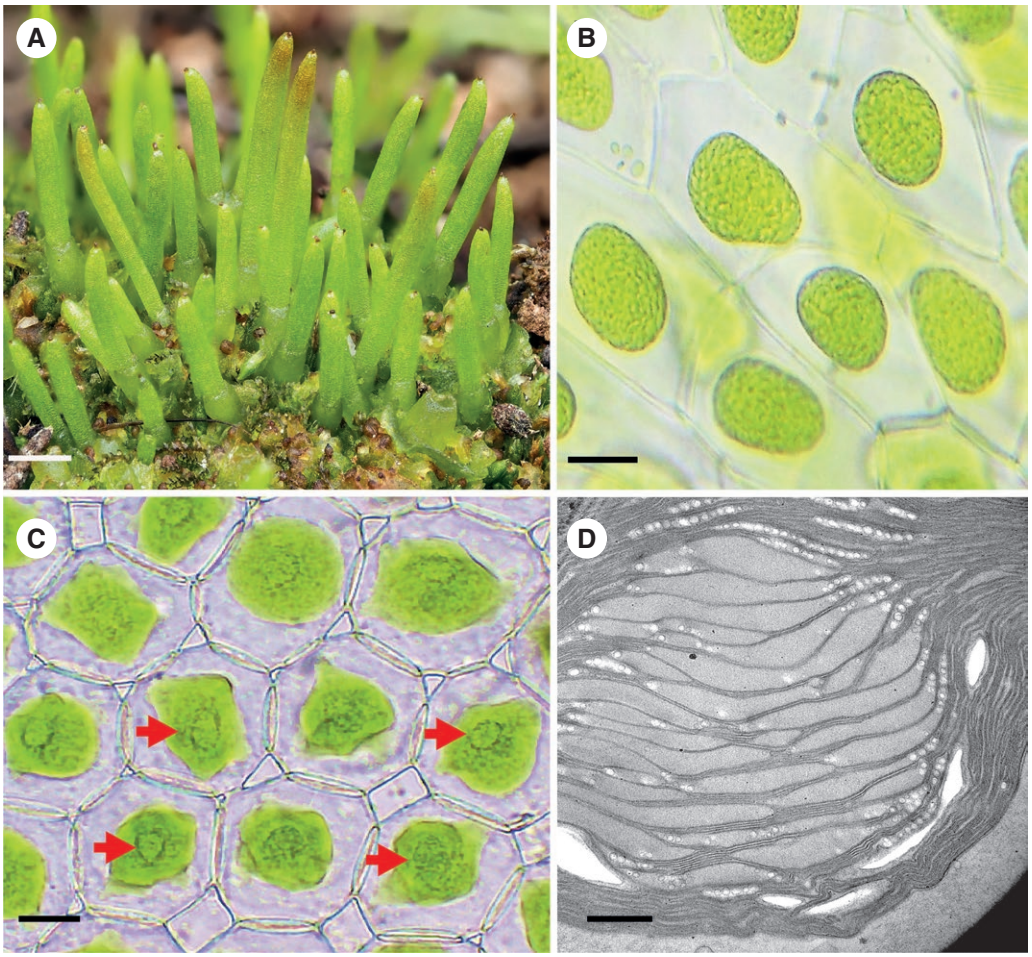

Fig. 3. Hornwort habit and plastid diversity. (A) Gametophyte and sporophyte of *Phaeoceros minutus* (Mitt.) S.W.Arnell (South Africa); (B) chloroplast without pyrenoid in *Nothoceros minarum* (Nees) J.C.Villarreal (Brazil); (C) chloroplast with pyrenoid in *Dendroceros crispatus* (Hook.) Nees (Colombia); (D) TEM image of the pyrenoid of *Phaeoceros carolinianus* (Michx.) Prosk. (USA). Scale bars: A, 2 mm; B–D, 10 µm. Pyrenoids are highlighted by red arrows. Photo credits: (A) Des Callaghan; (B–C) G.F.P.B.; and (D) K.S.R.

of the taxon may be due to the low rate of RNA editing in *Leiosporoceros* (Villarreal *et al.*, 2018). RNA edited sites contribute significantly to the overall number of phylogenetically informative positions, thereby influencing the reconstruction of phylogeny. A recent nuclear gene phylogeny by Bechteler *et al.* (2023) that included 12 species of hornworts also questioned the placement of *Leiosporoceros*, suggesting high levels of phylogenomic incongruence at the *Leiosporoceros* node.

We recovered two distinct topologies among hornworts, depending on the method of analysis. ML concatenation analysis strongly supports *Leiosporoceros* as sister to all hornworts (T1). However, concatenation implicitly assumes that all genes share the same history, and therefore it does not account for ILS or reticulation. This can introduce biases, which may lead to inconsistencies, lack of recombination or erroneous results with high confidence (Song *et al.*, 2012; Xi *et al.*, 2014; Cai *et al.*, 2021). Recently, Shen *et al.* (2024), using transcriptome data from nine hornworts, placed *Leiosporoceros* as the sister to all hornworts, aligning with our T1. However, the node values for *Leiosporoceros* (see phyto_22, table S7, and fig. S10 in Shen *et al.*, 2024) show that the proportion of gene duplications (GD ratio) and the number of ABAB-type duplications decrease slightly as support increases from 50 % (GD ratio: 6.13 %; ABAB 137: 43.35 %) to 80 % (GD ratio: 5.33 %; ABAB 109: 40.37 %). Additionally, the absence of ABAX duplications at both levels of support may suggest a lack of strong lineage separation at this node. This is consistent with the retention of some duplicated genes in both lineages, which could be related to patterns of ILS.

In contrast, our ASTRAL-III analyses, which are based on a multispecies coalescent (MSC) model, tend to recover *Leiosporoceros* sister to Anthocerotaceae (T2), although with low support (Song *et al.*, 2012; Xi *et al.*, 2014; Reddy *et al.*, 2017; Zhang *et al.*, 2018; Cai *et al.*, 2021). The ASTRAL results revealed gCF and sCF values were similar between T1 and T2 in the various datasets analysed, with high levels of discordance in the placement of *Leiosporoceros* (gCF and sCF < 37 %; see Fig. 1). The two topologies have nearly the same proportion of trees and sites supporting them, and the lack of resolution might be attributed to the presence of ILS. ILS is a prominent biological factor responsible for generating higher levels of inconsistency among topologies and has been reported in nearly all plant groups (Pérez-Escobar *et al.*, 2021). In liverworts, Marchantiales and Jungermanniopsida exhibit cytonuclear inconsistencies in their topologies due to the presence of ILS (Dong *et al.*, 2022), which also appears to occur in the moss genus *Sphagnum* L. (Meleshko *et al.*, 2021).

Evidence of ancient reticulation in hornworts

Our analyses confirm the presence of ancestral reticulation and introgressions among the major ancestral lineages of hornworts (Fig. 2), suggesting complex patterns of gene flow among genera during their diversification. This contributes to the discordance observed in T1 and T2 (Figs 1 and 2) as well as the phylogenetic position of *Leiosporoceros*. Hybridization has proven to be a significant source of variation in the evolutionary history of organisms (Anderson, 1949; Taylor and Larson, 2019) because of the potential to generate novel characteristics more rapidly than other processes such as mutation (Soltis, 2013; Suarez-Gonzalez *et al.*, 2018; Stull *et al.*, 2023).

Although hybridization has been suggested in hornworts (Proskauer, 1957, 1969), it has been considered to be rare or of limited evolutionary importance in the group. In contrast, hybrids have been reported between species in mosses and liverworts (Natcheva and Cronberg, 2004) and ancient reticulation has been proposed between major liverwort classes (Dong *et al.*, 2022). For example, in mosses, there is evidence of hybridization between ancestral lineages of *Sphagnum* (Natcheva and Cronberg, 2007). Consequently, recent gene flow between species within the genus is limited, despite the interspecific hybridization documented in the group (Meleshko *et al.*, 2021). Similarly, in liverworts, signals of ancient hybridization and ILS were reported in the backbone phylogeny involving the ancestors of simple thalloids, leafy liverworts and complex thalloids (Dong *et al.*, 2022). Introgression and hybridization events have been identified in groups across the tree of life, including cyanobacteria (Pardo-De la Hoz *et al.*, 2023), lichens (Keuler *et al.*, 2022), ferns (Huang *et al.*, 2020; Chen *et al.*, 2023), angiosperms (Sun *et al.*, 2020; Debray *et al.*, 2022; Hodel *et al.*, 2022), gymnosperms (Liu *et al.*, 2022), dragonflies (Suvorov *et al.*, 2022), snakes (Schöneberg *et al.*, 2023) and primates (Vanderpool *et al.*, 2020). The widespread occurrence of these events underscores the fundamental role of hybridization in shaping the evolutionary trajectory and diversification of organisms.

Timing of diversification in hornworts and the evolution of the pyrenoid

Since the placement of *Leiosporoceros* remains elusive, we explored the timing of hornwort diversification in two distinct scenarios (T1 and T2). Our estimates place the origin of the crown age of hornworts in the Carboniferous period, in line with previous studies (Villarreal and Renner, 2012; Morris *et al.*, 2018; Harris *et al.*, 2022; Bechteler *et al.*, 2023; Shen *et al.*, 2024). The Carboniferous and Cretaceous–Palaeogene timing of hornwort diversification may have been influenced by the environmental characteristics of the time, especially significantly low CO_2 levels (Fig. 1). A dramatic drop in atmospheric CO_2 levels, starting at >275 ppm in the Devonian and down to <190 ppm in the Carboniferous (Igamberdiev and Lea, 2006), was speculated to trigger the evolution of pyrenoid carbon-concentrating mechanisms independently in 'algae': including euglenophytes, glaucocystophytes, straminopiles, dinoflagellates, and green and red algal lineages (Badger *et al.*, 2000). The convergent evolution of pyrenoids is thought to have occurred in aquatic environments to overcome the low diffusion of CO_2 in water, which was particularly problematic with the dramatic drop in atmospheric CO_2 levels. The evolution of the hornwort pyrenoid has remained enigmatic because of the terrestrial nature of these plants.

The topology in which *Leisporoceros* is sister to the rest of the hornworts (T1), and the absence of pyrenoids in *Leiosporoceros* and outgroup taxa (other land plants) weighed heavily on previous ancestral trait reconstructions that identified a *de novo* origin of the hornwort pyrenoid (Villarreal and Renner, 2012). It is possible that pyrenoids in hornworts evolved in response

to the dramatic drop of CO_2 in the Carboniferous as in algal groups (Villarreal and Renner, 2012; He *et al.*, 2023; Ruaud *et al.*, 2024). The alternative topology presented here (T2) suggests that the ancestor of the hornworts may have possessed the trait. Since the pyrenoid is widespread in charophyte lineages, we cannot rule out the possibility that the pyrenoid was present in the ancestor of all land plants and subsequently lost in setaphytes (mosses and liverworts) and tracheophytes (Robison *et al.*, 2025). Following this interpretation, pyrenoids would have disappeared several times during hornwort evolution, including in *Leiosporoceros*, *Phaeomegaceros*, *Megaceros*, *Nothoceros* and sporadically in species-rich genera with mostly pyrenoid-containing taxa such as *Anthoceros* and *Folioceros*.

Atavistic gains of the pyrenoid are also indicated in both interpretations and these confound our understanding of the evolutionary history of the hornwort pyrenoid (Fig. 3). For example, the occurrence of pyrenoids varies across populations even within a single species of *Nothoceros vincentianus* (Lehm. et Lindenb.) J.C.Villarreal in which pyrenoidless and pyrenoid-containing plants have been reported while the remaining species in the genus are devoid of pyrenoids (Villarreal and Renner, 2014). Similarly, the structurally distinct and complex pyrenoid in *Dendroceros* may be interpreted as a reappearance associated with the unique microenvironment of this epiphytic genus as the pyrenoid has disappeared from the vast majority of its sister taxa in the Dendrocerotaceae (Vaughn *et al.*, 1992; Schuette and Renzaglia, 2010; Villarreal and Renner, 2012). Although our analyses do not put to rest the debate over the origin of the hornwort pyrenoid, they do support the viability of duelling interpretations that pyrenoids were plesiomorphic and lost repeatedly within hornworts or that they evolved in response to lower CO_2 levels during hornwort diversification. Because pyrenoids are not that readily visible, especially in dried specimens or if abundant starch is present, there is a void of data on their occurrence across taxa worldwide. A clear understanding of the evolution of the hornwort pyrenoid awaits more comprehensive studies using sophisticated methods such as high-resolution and cryo-microscopy, and the elucidation of the development, genetic underpinnings and physiology of the photosynthetic apparatus in hornworts Robison *et al.*, (2025).

Our findings also highlight a period of heightened diversification in hornworts during the mid-Cretaceous and Palaeogene with nearly synchronous emergence of all genera, similar to the cycad radiation (Nagalingum *et al.*, 2011; Condamine *et al.*, 2015). This period of increased diversification aligns with the aftermath of the Cretaceous–Palaeogene (K–Pg) mass extinction (Springer *et al.*, 2003; Renne *et al.*, 2013; Vajda and Bercovici, 2014). Changes in pollen and spore assemblages across the K–Pg boundary have provided crucial insights into vegetation responses during a global environmental crisis triggered by the asteroid impact (Vajda and Bercovici, 2014). The Cretaceous period is also characterized by significant shifts in terrestrial plant and animal communities, such as the radiation of flowering plants (Davis *et al.*, 2005; Lutzoni *et al.*, 2018; Donoghue *et al.*, 2021) and mammal diversification (Vajda and Bercovici, 2014; Ramírez-Barahona *et al.*, 2020).

Setaphytes also underwent significant diversification during the Cretaceous–Palaeogene period. For example, species-rich lineages of liverworts and mosses, such as Hypnales W.R.Buck & Vitt and Lejeuneaceae Rostovzev, diversified nearly in parallel with angiosperms during this time (Feldberg *et al.*, 2014; Laenen *et al.*, 2014; Bechteler *et al.*, 2023; Shen *et al.*, 2024). Additionally, there is evidence of profound divergences in complex thalloid liverwort species near the K/T global extinction event. These species exhibited a xerophytic lifestyle with desiccation tolerance strategies (e.g. *Corsinia* Raddi, *Exormotheca* Mitt., *Plagiochasma* Lehm., *Riccia* L., *Targionia* L.; Bischler-causse *et al.*, 2012; Villarreal *et al.*, 2016) important to adaptations in a changing terrestrial landscape. Hornworts are now added to the mix of a wide range of land plants that experienced bursts of diversification during this period, including setaphytes, leptosporangiate ferns (Schuettpelz and Pryer, 2009), conifers (Taxodiaceae, Cupressaceae, Araucariaceae; Axsmith and Jacobs, 2005), cycads, *Ginkgo* (Johnson, 2002; Vajda and Bercovici, 2014) and angiosperms (Davis *et al.*, 2005; Lutzoni *et al.*, 2018; Donoghue *et al.*, 2021).

SUPPLEMENTARY DATA

Supplementary data are available at *Annals of Botany* online and consist of the following.

Figure S1: Hornwort phylogeny inferred with the concatenation method (IQ-TREE) from nucleotide data showing bootstrap support with 234 genes. **Figure S2:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data including the local posterior probability (Lpp) values with 234 genes. **Figure S3:** Hornwort phylogeny inferred with the concatenation method (IQ-TREE) from nucleotide data with 195 genes (only genes in *Leiosporoceros dussii*), including bootstrap support and concordance values for each node outlined in Table S4. **Figure S4:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data with 195 genes (only genes in *Leiosporoceros dussii*), including local posterior probability and concordance values for each node outlined in Table S4. **Figure S5:** Hornwort phylogeny inferred with the concatenation method (IQ-TREE) from nucleotide data with 234 genes including bootstrap support and concordance values for each node outlined in Table S4. **Figure S6:** Hornwort phylogeny inferred with the concatenation method (IQ-TREE) from nucleotide data with 161 consistent genes including bootstrap support and concordance values for each node outlined in Table S4. **Figure S7:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data with 234 genes including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S8:** Hornwort phylogeny inferred from nucleotide data with 234 genes including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S9:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data with 161 consistent genes including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S10:** Hornwort phylogeny inferred from nucleotide data with 161 consistent genes including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S11:** Hornwort phylogeny inferred from nucleotide data with 195 genes (only genes in *Leiosporoceros*

dussii), including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S12:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data with 133 consistent genes (only genes in *Leiosporoceros dussii*), including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S13:** Hornwort phylogeny inferred from nucleotide data with 133 consistent genes (only genes in *Leiosporoceros dussii*), including local posterior probability values and concordance values for each node outlined in Table S4. **Figure S14:** Estimates of ancient reticulate evolution (SNaQ) from all hornwort genera using 234 genes and $h_{max} = 3$ (number of hybridizations). **Figure S15:** Estimates of ancient reticulate evolution (SNaQ) from all hornwort genera using 234 genes and $h_{max} = 3$ (number of hybridizations). **Figure S16:** Plot scores between hybrid values. Values of h_{max} are the number of hybridization events (0–3). It was based on topology T1 (*Leiosporoceros* sister to all hornworts). **Figure S17:** Plot scores between hybrid values. Values of h_{max} are the number of hybridization events (0–3). It was based on topology T2 (*Leiosporoceros* sister to Anthocerotaceae). **Figure S18:** Divergence time estimates for hornwort genera using the concatenation method (IQ-TREE) based on 234 nuclear genes. **Figure S19:** Divergence time estimates for hornwort genera using the coalescence-based method (ASTRAL-III) based on 234 nuclear genes. **Figure S20:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data including the quartet values with 234 genes, for the three possible topologies, presented as percentages near the node (100/0/0). **Figure S21:** Hornwort phylogeny inferred with the coalescence-based method (ASTRAL-III) from nucleotide data including the local posterior probability (lpp) values with 234 genes, for the three possible topologies, presented as percentages near the node (100/0/0). **Table S1:** Taxon sampling indicating accession numbers, DNA sample IDs and respective voucher information (i.e. Collection number, Locality and Herbarium). **Table S2:** Consistent and inconsistent genes between concatenation-based IQ-TREE (T1) and quartet-based ASTRAL (T2) in hornworts. **Table S3:** Concordance factors for the nucleotide data stemming from IQ-TREE and ASTRAL analyses from hornworts. **Table S4:** Test of alternative topologies in IQ-TREE, identifying the most influential genes with partition-wise log likelihoods. **Table S5:** Analyses with SNaQ to explore possible reticulation events in the evolutionary history of hornworts. **Table S6:** Analysis in HyDe with gene flow for hornworts. Database with 234 genes and all hornwort genera.

FUNDING

This work was supported by the NSF collaborative project 'Building a Comprehensive Evolutionary History of Flagellate Plants' [DEB 1541506 to J.G.B., E.C.D., S.F.M. and E.B.S.]. P.S. was funded by the NSF IOS [2109789]. G.F.P.B. was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior, Brazil [CAPES – Finance Code 001], the Latin American and the Caribbean Macro Universities, the postgraduate Mobility Program 02-2017, the International Association for Plant Taxonomy [IAPT – Research Grant 2019], the Emerging Leaders in the Americas Program – ELAP [2018-2019] and the Rufford Foundation [Small Grants 2018-2019]. J.C.V.A. was funded by the Canada Research Chair [950-232698] the CRNSG- RGPIN [05967–2016] and the Canadian Foundation for Innovation [projects 36781, 39135].

ACKNOWLEDGMENTS

We thank all the herbaria that provided specimens and the anonymous reviewers and the Associate Editor for their constructive comments on previous versions of the manuscript. G.F.P.B. and J.C.V.A. thank Adriel Sierra and the Louis-Marie Herbarium (QFA) for support.

AUTHOR CONTRIBUTIONS

The concept and methodology of the study were developed by G.F.B.P., J.G.B., K.S.R. and J.C.V.A.; specimens were collected by G.F.B.P., J.G.B., K.R., A.S.M.S., D.C.C., D.B., E.B.S., F.W.L., S.F.M., E.C.D., L.E., N.S.A, P.S., S.C., J.G.D., S.P., C.S.L. and J.C.V.A.; DNA was extracted and data were generated by G.F.B.P., J.G.B., L.E. and J.C.V.A.; analyses were performed by G.F.B.P. and J.C.V.A.; writing – original draft preparation by G.F.B.P., K.S.R. and J.C.V.A.; writing – review and editing by G.F.B.P., J.G.B., K.S.R., A.S.M.S., D.C.C, D.B., E.B.S., F.W.L., S.F.M., E.C.D., L.E., N.S.A, P.S., S.C., J.G.D., S.P., C.S.L. and J.C.V.A.; funding for the study was obtained by G.F.B.P., J.G.B., E.C.D., E.B.S., S.F.M., P.S. and J.C.V.A. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AVAILABILITY OF DATA

The raw sequence reads for all samples have been deposited in the NCBI SRA database (and all SRA accession numbers can be found in Supplementary Data Table S1). All scripts, phylogenetic nucleotide alignments, and resulting gene trees and species trees are available at Github: https://github.com/gpenalozabojaca/Hornwort-diversification-.git.

LITERATURE CITED

Anderson E. 1949. *Introgressive hybridization.* New York: J. Wiley.

Archangelsky S, **Villar de Seone L. 1996**. Estudios palinógicos de la formación Baqueró (Cretácico), provincia de Santa Cruz, Argentina. *Ameghiniana* **33**: 7–19.

Atkinson N, **Feike D**, **Mackinder L**, ***et al***. **2016**. Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. *Plant Biotechnology Journal* **14**: 1302–1315.

Axsmith BJ, **Jacobs BF. 2005**. The conifer *Frenelopsis ramosissima* (Cheirolepidiaceae) in the Lower Cretaceous of Texas: systematic, biogeographical, and paleoecological implications. *International Journal of Plant Sciences* **166**: 327–337.

Badger MR, **von Caemmerer S**, **Ruuska S**, **Nakano H. 2000**. Electron flow to oxygen in higher plants and algae: rates and control of direct photoreduction (Mehler reaction) and rubisco oxygenase. *Philosophical*

Transactions of the Royal Society of London, Series B: Biological Sciences **355**: 1433–1446.

Badger MR, **Hanson D**, **Price GD. 2002**. Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Functional Plant Biology* **29**: 161–173.

Bechteler J, **Peñaloza-Bojacá G**, **Bell D**, ***et al***. **2023**. Comprehensive phylogenomic time tree of bryophytes reveals deep relationships and uncovers gene incongruences in the last 500 million years of diversification. *American Journal of Botany* **110**: 1–20.

Bell D, **Lin Q**, **Gerelle WK**, ***et al***. **2020**. Organellomic data sets confirm a cryptic consensus on (unrooted) land-plant relationships and provide new insights into bryophyte molecular evolution. *American Journal of Botany* **107**: 91–115.

Berner RA. 2001. GEOCARB III: A revised model of atmospheric CO_2 over Phanerozoic time. *American Journal of Science* **301**: 182–204.

Bischler-causse H, **Gradstein R**, **Jovet-ast S**, **Long D**, **Allen NS. 2012**. Marchantiidae. *Flora Neotropica Monographs* **97**: 1–267.

Blischak PD, **Chifman J**, **Wolfe AD**, **Kubatko LS. 2018**. HyDe: a Python package for genome-scale hybridization detection. *Systematic Biology* **67**: 821–829.

Breinholt JW, **Carey SB**, **Tiley GP**, ***et al***. **2021**. A target enrichment probe set for resolving the flagellate land plant tree of life. *Applications in Plant Sciences* **9**: 1–28.

Brown J, **Walker J**, **Smith S. 2017**. Phyx: phylogenetic tools for unix. *Bioinformatics* **33**: 1886–1888.

Cai L, **Xi Z**, **Lemmon EM**, ***et al***. **2021**. The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology* **70**: 491–507.

Chan KO, **Hutter CR**, **Wood PL**, **Grismer LL**, **Brown RM. 2020**. Target-capture phylogenomics provide insights on gene and species tree discordances in Old World treefrogs (Anura: Rhacophoridae). *Proceedings Biological Sciences* **287**: 20202102.

Chen H, **Fang Y**, **Zwaenepoel A**, **Huang S**, **Van de Peer Y**, **Li Z. 2023**. Revisiting ancient polyploidy in Leptosporangiate ferns. *The New Phytologist* **237**: 1405–1417.

Chitaley S, **Yawale N. 1980**. On *Notothylites nirulai* gen. et sp. nov. a petrified sporogonium from the Deccan-Intertrappean beds of Mohgaonkalan M. P. (India). *Botanique* **9**: 111–118.

Condamine FL, **Nagalingum NS**, **Marshall CR**, **Morlon H. 2015**. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evolutionary Biology* **15**: 65.

Cooper BJ, **Moore MJ**, **Douglas NA**, ***et al***. **2023**. Target enrichment and extensive population sampling help untangle the recent, rapid radiation of *Oenothera* sect. *Calylophus*. *Systematic Biology* **72**: 249–263.

Davis CC, **Webb CO**, **Wurdack KJ**, **Jaramillo CA**, **Donoghue MJ. 2005**. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. *The American Naturalist* **165**: E36–E65.

Debray K, **Le Paslier M-C**, **Bérard A**, ***et al***. **2022**. Unveiling the patterns of reticulated evolutionary processes with phylogenomics: hybridization and polyploidy in the genus *Rosa*. *Systematic Biology* **71**: 547–569.

Degnan JH, **Rosenberg NA. 2009**. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* **24**: 332–340.

Dong S, **Yu J**, **Zhang L**, **Goffinet B**, **Liu Y. 2022**. Phylotranscriptomics of liverworts: revisiting the backbone phylogeny and ancestral gene duplications. *Annals of Botany* **130**: 951–964.

Donoghue P, **Harrison C**, **Paps J**, **Schneider H. 2021**. The evolutionary emergence of land plants. *Current Biology* **31**: R1281–R1298.

Doyle J, **Doyle J. 1987**. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Duff J, **Villarreal JC**, **Cargill C**, **Renzaglia K. 2007**. Progress and challenges toward developing a phylogeny and classification of the hornworts. *The Bryologist* **110**: 214–243.

Feldberg K, **Schneider H**, **Stadler T**, **Schäfer-Verwimp A**, **Schmidt AR**, **Heinrichs J. 2014**. Epiphytic leafy liverworts diversified in angiosperm-dominated forests. *Scientific Reports* **4**: 5974.

Feldberg K, **Gradstein SR**, **Gröhn C**, ***et al***. **2021**. Checklist of fossil liverworts suitable for calibrating phylogenetic reconstructions. *Bryophyte Diversity and Evolution* **43**: 14–71.

Graham A. 1987. Miocene communities and paleoenvironments of southern Costa Rica. *American Journal of Botany* **74**: 1501–1518.

Guan Y, **Liu L**, **Wang Q**, ***et al***. **2018**. Gene refashioning through innovative shifting of reading frames in mosses. *Nature Communications* **9**: 1555.

Harris BJ, **Clark JW**, **Schrempf D**, ***et al***. **2022**. Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. *Nature Ecology & Evolution* **6**: 1634–1643.

Hasegawa J. 1988. A proposal for a new system of the Anthocerotae, with a revision of the genera. *Journal of the Hattori Botanical Laboratory* **64**: 87–95.

Hässel de Menéndez G. 1988. A proposal for a new classification of the genera within the Anthocerotophyta. *Journal of the Hattori Botanical Laboratory* **64**: 71–86.

He S, **Crans VL**, **Jonikas MC. 2023**. The pyrenoid: the eukaryotic CO_2-concentrating organelle. *The Plant Cell* **35**: 3236–3259.

Heureux A, **Young J**, **Whitney S**, ***et al***. **2017**. The role of Rubisco kinetics and pyrenoid morphology in shaping the CCM of haptophyte microalgae. *Journal of Experimental Botany* **68**: 3959–3969.

Hoang DT, **Chernomor O**, **von Haeseler A**, **Minh BQ**, **Vinh L. 2017**. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* **35**: 518–522.

Hodel RGJ, **Zimmer EA**, **Liu B-B**, **Wen J. 2022**. Synthesis of nuclear and chloroplast data combined with network analyses supports the polyploid origin of the apple tribe and the hybrid origin of the Maleae—Gillenieae clade. *Frontiers in Plant Science* **12**: 1–15.

Huang C, **Qi X**, **Chen D**, **Qi J**, **Ma H. 2020**. Recurrent genome duplication events likely contributed to both the ancient and recent rise of ferns. *Journal of Integrative Plant Biology* **62**: 433–455.

Igamberdiev AU, **Lea PJ. 2006**. Land plants equilibrate O_2 and CO_2 concentrations in the atmosphere. *Photosynthesis Research* **87**: 177–194.

Ignatov MS, **Maslova EV. 2021**. Fossil mosses: What do they tell us about moss evolution? *Bryophyte Diversity and Evolution* **43**: 72–97.

Johnson KR. 2002. Megaflora of the Hell Creek and Lower Fort Union formations in the Western Dakotas: vegetational response to climate change, the cretaceous-tertiary boundary event, and rapid marine transgression. In: *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: An Integrated continental record of the end of the Cretaceous.* Boulder, CO: Geological Society of America, 329–391.

Kalyaanamoorthy S, **Minh BQ**, **Wong TKF**, **von Haeseler A**, **Jermiin LS. 2017**. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**: 587–589.

Keuler R, **Jensen J**, **Barcena-Peña A**, ***et al***. **2022**. Interpreting phylogenetic conflict: hybridization in the most speciose genus of lichen-forming fungi. *Molecular Phylogenetics and Evolution* **174**: 107543.

Kishino H, **Hasegawa M. 1989**. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *Journal of Molecular Evolution* **29**: 170–179.

Kishino H, **Miyata T**, **Hasegawa M. 1990**. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *Journal of Molecular Evolution* **31**: 151–160.

Knowles LL, **Huang H**, **Sukumaran J**, **Smith SA. 2018**. A matter of phylogenetic scale: Distinguishing incomplete lineage sorting from lateral gene transfer as the cause of gene tree discord in recent versus deep diversification histories. *American Journal of Botany* **105**: 376–384.

Laenen B, **Shaw B**, **Schneider H**, ***et al***. **2014**. Extant diversity of bryophytes emerged from successive post-Mesozoic diversification bursts. *Nature Communications* **5**: 5134.

Lanfear R, **Minh B**, **Hahn M. 2018**. *Calculating and interpreting gene- and site-concordance factors in phylogenomics.* http://www.robertlanfear.com/blog/files/concordance_factors.html (16 May 2022).

Leebens-Mack J, **Barker M**, **Carpenter E**, ***et al***. **2019**. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679–685.

Li F-W, **Villarreal JC**, **Szövényi P. 2017**. Hornworts: an overlooked window into carbon-concentrating mechanisms. *Trends in Plant Science* **22**: 275–277.

Liu X-Q, **Xia X-M**, **Chen L**, **Wang X-Q. 2022**. Phylogeny and evolution of Cupressaceae: updates on intergeneric relationships and new insights on ancient intergeneric hybridization. *Molecular Phylogenetics and Evolution* **177**: 107606.

Lutzoni F, **Nowak MD**, **Alfaro ME**, ***et al***. **2018**. Contemporaneous radiations of fungi and plants linked to symbiosis. *Nature Communications* **9**: 5451.

Meleshko O, **Martin MD**, **Korneliussen TS**, ***et al***. **2021**. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not

ongoing introgression in a rapidly radiated bryophyte genus. *Molecular Biology and Evolution* **38**: 2750–2766.
Meyer M, **Whittaker C**, **Griffiths H. 2017**. The algal pyrenoid: key unanswered questions. *Journal of Experimental Botany* **68**: 3739–3749.
Meyer M, **Goudet MMM**, **Griffiths H. 2020**. The algal pyrenoid. In: **Larkum A**, **Grossmann A**, **Raven J**. eds. *Photosynthesis in algae: biochemical and physiological mechanisms*, Vol. **45**. Berlin: Springer Nature, 179–203.
Minh B, **Trifinopoulos J**, **Schrempf D**, **Schmidt HA. 2018**. *IQ-TREE version 2.2.0: Tutorials and Manual Phylogenomic software by maximum likelihood*: 1–156. http://www.iqtree.org/doc/iqtree-doc.pdf (25 May 2023).
Minh BQ, **Trifinopoulos J**, **Schrempf D**, **Schmidt HA. 2019**. *IQ-TREE version 2.0: Tutorials and manual phylogenomic software by maximum likelihood. Manual/tutorial*. https://github.com/iqtree (30 May 2023).
Minh BQ, **Hahn MW**, **Lanfear R. 2020**. New methods to calculate concordance factors for phylogenomic datasets. *Molecular Biology and Evolution* **37**: 2727–2733.
Moisan P, **Voigt S**, **Schneider JW**, **Kerp H. 2012**. New fossil bryophytes from the Triassic Madygen Lagerstätte (SW Kyrgyzstan). *Review of Palaeobotany and Palynology* **187**: 29–37.
Morris JL, **Puttick MN**, **Clark JW**, ***et al***. **2018**. The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* **115**: E2274–E2283.
Nagalingum NS, **Marshall CR**, **Quental TB**, **Rai HS**, **Little DP**, **Mathews S. 2011**. Recent synchronous radiation of a living fossil. *Science* **334**: 796–799.
Natcheva R, **Cronberg N. 2004**. What do we know about hybridization among bryophytes in nature? *Canadian Journal of Botany* **82**: 1687–1704.
Natcheva R, **Cronberg N. 2007**. Maternal transmission of cytoplasmic DNA in interspecific hybrids of peat mosses, *Sphagnum* (Bryophyta). *Journal of Evolutionary Biology* **20**: 1613–1616.
Nguyen L-T, **Schmidt H**, **von Haeseler A**, **Minh BQ. 2015**. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**: 268–274.
Pardo-De la Hoz CJ, **Magain N**, **Piatkowski B**, ***et al***. **2023**. Ancient rapid radiation explains most conflicts among gene trees and well-supported phylogenomic trees of Nostocalean Cyanobacteria. *Systematic Biology* **72**: 694–712.
Peñaloza-Bojacá GF, **de Oliveira MF**, **Maciel-Silva AS. 2020**. *Anthoceros tristanianus* J.C. Villarreal, J.J. Engel & Váňa in Váňa & Engel (Anthocerotaceae, Anthocerotophyta) a new record for South America from the Colombian bryophyte flora. *Cryptogamie Bryologie* **41**: 105.
Pérez-Escobar OA, **Dodsworth S**, **Bogarín D**, ***et al***. **2021**. Hundreds of nuclear and plastid loci yield novel insights into orchid relationships. *American Journal of Botany* **108**: 1166–1180.
Proskauer J. 1957. Studies on Anthocerotales V. *Phytomorphology* **7**: 113–135.
Proskauer J. 1969. Studies on Anthocerotales. VIII. *Phytomorphology* **19**: 52–66.
R Core Team. 2024. *R: a language and environment for statistical computing*. https://www.r-project.org/ (4 December 2024).
Rambaut A. 2017. *FigTree-version 1.4. 3, a graphical viewer of phylogenetic trees*. http://tree.bio.ed.ac.uk/software/figtree/ (20 August 2023).
Ramírez-Barahona S, **Sauquet H**, **Magallón S. 2020**. The delayed and geographically heterogeneous diversification of flowering plant families. *Nature Ecology & Evolution* **4**: 1232–1238.
Reddy S, **Kimball RT**, **Pandey A**, ***et al***. **2017**. Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. *Systematic Biology* **66**: 857–879.
Renne P, **Deino A**, **Hilgen F**, ***et al***. **2013**. Time scales of critical events around the Cretaceous-Paleogene boundary. *Science* **339**: 684–687.
Renzaglia K, **Schuette S**, **Duff J**, ***et al***. **2007**. Bryophyte phylogeny: advancing the molecular and morphological frontiers. *The Bryologist* **110**: 179–213.
Renzaglia K, **Villarreal JC**, **Duff RJ**, **Goffinet B. 2009**. New insights into morphology, anatomy, and systematics of hornworts In: **Shaw AJ**. ed. *Bryophyte biology*. Cambridge: Cambridge University Press, 139–172.
Ruaud S, **Nötzold SI**, **Waller M**, ***et al***. **2024**. Molecular underpinnings of hornwort carbon concentrating mechanisms: subcellular localization of putative key molecular components in the model hornwort *Anthoceros agrestis*. *bioRxiv*: 1–36.
Sayyari E, **Mirarab S. 2016**. Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution* **33**: 1654–1668.
Schöneberg Y, **Winter S**, **Arribas O**, ***et al***. **2023**. Genomics reveals broad hybridization in deeply divergent Palearctic grass and water snakes (*Natrix* spp.). *Molecular Phylogenetics and Evolution* **184**: 107787.
Schuette S, **Renzaglia K. 2010**. Development of multicellular spores in the hornwort genus *Dendroceros* (Dendrocerotaceae, Anthocerotophyta) and the occurrence of endospory in Bryophytes. *Nova Hedwigia* **91**: 301–316.
Schuettpelz E, **Pryer KM. 2009**. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 11200–11205.
Shen X-X, **Steenwyk JL**, **Rokas A. 2021**. Dissecting incongruence between concatenation- and quartet-based approaches in phylogenomic data. *Systematic Biology* **70**: 997–1014.
Shen C, **Li H**, **Shu L**, **Huang W**, **Zhu R. 2024**. Ancient large-scale gene duplications and diversification in bryophytes illuminate the plant terrestrialization. *New Phytologist* **1**: 1–17.
Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* **51**: 492–508.
Shimodaira H, **Hasegawa M. 1999**. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16**: 1114–1116.
Smith S, **O'Meara B. 2012**. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**: 2689–2690.
Söderström L, **Hagborg A**, **von Konrat M**, ***et al***. **2016**. World checklist of hornworts and liverworts. *PhytoKeys* **59**: 1–828.
Solís-Lemus C, **Ané C. 2016**. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics* **12**: e1005896.
Solís-Lemus C, **Bastide P**, **Ané C. 2017**. PhyloNetworks: a package for phylogenetic networks. *Molecular Biology and Evolution* **34**: 3292–3298.
Soltis PS. 2013. Hybridization, speciation and novelty. *Journal of Evolutionary Biology* **26**: 291–293.
Song S, **Liu L**, **Edwards SV**, **Wu S. 2012**. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 14942–14947.
Springer MS, **Murphy WJ**, **Eizirik E**, **O'Brien SJ. 2003**. Placental mammal diversification and the Cretaceous–Tertiary boundary. *Proceedings of the National Academy of Sciences* **100**: 1056–1061.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313.
Steinthorsdottir M, **Vajda V. 2015**. Early Jurassic (late Pliensbachian) CO_2 concentrations based on stomatal analysis of fossil conifer leaves from eastern Australia. *Gondwana Research* **27**: 932–939.
Strimmer K, **Rambaut A. 2002**. Inferring confidence sets of possibly misspecified gene trees. *Proceedings Biological Sciences* **269**: 137–142.
Stull GW, **Pham KK**, **Soltis PS**, **Soltis DE. 2023**. Deep reticulation: the long legacy of hybridization in vascular plant evolution. *The Plant Journal: for Cell and Molecular Biology* **114**: 743–766.
Suarez-Gonzalez A, **Lexer C**, **Cronk QCB. 2018**. Adaptive introgression: a plant perspective. *Biology Letters* **14**: 20170688.
Sun M, **Folk R**, **Gitzendanner M**, ***et al***. **2020**. Recent accelerated diversification in Rosids occurred outside the tropics. *Nature Communications* **11**: 3333.
Suvorov A, **Scornavacca C**, **Fujimoto MS**, ***et al***. **2022**. Deep ancestral introgression shapes evolutionary history of dragonflies and damselflies. *Systematic Biology* **71**: 526–546.
Taylor SA, **Larson EL. 2019**. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution* **3**: 170–177.
Uniacke J, **Zerges W. 2009**. Chloroplast protein targeting involves localized translation in *Chlamydomonas*. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 1439–1444.
Vajda V, **Bercovici A. 2014**. The global vegetation pattern across the Cretaceous–Paleogene mass extinction interval: a template for other extinction events. *Global and Planetary Change* **122**: 29–49.
Vanderpool D, **Minh BQ**, **Lanfear R**, ***et al***. **2020**. Primate phylogenomics uncovers multiple rapid radiations and ancient interspecific introgression. *PLoS Biology* **18**: e3000954.
Vaughn K, **Ligrone R**, **Owen H**, ***et al***. **1992**. The anthocerote chloroplast: a review. *New Phytologist* **120**: 169–190.
Villarreal JC, **Renner S. 2012**. Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proceedings of the National Academy of Sciences* **109**: 18873–18878.

Villarreal JC, **Renner S. 2013**. Correlates of monoicy and dioicy in hornworts, the apparent sister group to vascular plants. *BMC Evolutionary Biology* **13**: 1–8.

Villarreal JC, **Renner S. 2014**. A review of molecular-clock calibrations and substitution rates in liverworts, mosses, and hornworts, and a timeframe for a taxonomically cleaned-up genus *Nothoceros*. *Molecular Phylogenetics and Evolution* **78**: 25–35.

Villarreal JC, **Renzaglia K. 2006**. Structure and development of *Nostoc* strands in *Leiosporoceros dussii* (Anthocerotophyta): a novel symbiosis in land plants. *American Journal of Botany* **93**: 693–705.

Villarreal JC, **Renzaglia K. 2015**. The hornworts: important advancements in early land plant evolution. *Journal of Bryology* **37**: 157–170.

Villarreal JC, **Cusimano N**, **Renner S. 2015**. Biogeography and diversification rates in hornworts: The limitations of diversification modeling. *Taxon* **64**: 229–238.

Villarreal JC, **Crandall-Stotler B**, **Hart M**, **Long D**, **Forrest L. 2016**. Divergence times and the evolution of morphological complexity in an early land plant lineage (Marchantiopsida) with a slow molecular rate. *New Phytologist* **209**: 1734–1746.

Villarreal JC, **Turmel M**, **Bourgouin-Couture M**, ***et al***. **2018**. Genome-wide organellar analyses from the hornwort *Leiosporoceros dussii* show low frequency of RNA editing. *PLoS One* **13**: e0200491.

Xi Z, **Liu L**, **Rest JS**, **Davis CC. 2014**. Coalescent versus concatenation methods and the placement of *Amborella* as sister to water lilies. *Systematic Biology* **63**: 919–932.

Zhan Y, **Marchand CH**, **Maes A**, ***et al***. **2018**. Pyrenoid functions revealed by proteomics in *Chlamydomonas reinhardtii*. *PLoS One* **13**: e0185039–e0185020.

Zhang C, **Rabiee M**, **Sayyari E**, **Mirarab S. 2018**. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**: 153.

Robison TA, **Oh ZG**, **Lafferty D**, ***et al***. **2025**. Hornworts reveal a spatial model for pyrenoid-based CO2-concentrating mechanisms in land plants. *Nature Plants* **11**: 63–73. doi:10.1038/s41477-024-01871-0.