



## Article

# Plastid Phylogenomics and Plastomic Diversity of the Extant Lycophytes

Sisi Chen <sup>1,2,3</sup>, Ting Wang <sup>1,4</sup>, Jiangping Shu <sup>1,3,5</sup> , Qiaoping Xiang <sup>2</sup>, Tuo Yang <sup>1</sup>, Xianchun Zhang <sup>2,\*</sup>  and Yuehong Yan <sup>1,\*</sup> 

- <sup>1</sup> Shenzhen Key Laboratory for Orchid Conservation and Utilization, National Orchid Conservation Center of China and the Orchid Conservation & Research Center of Shenzhen, Shenzhen 518114, China; 17770972237@163.com (S.C.); tingwang93@126.com (T.W.); jpshu@scbg.ac.cn (J.S.); yangtuo416@outlook.com (T.Y.)
- <sup>2</sup> State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Science, Beijing 100093, China; qpxiang@ibcas.ac.cn
- <sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China
- <sup>4</sup> College of Biodiversity Conservation, Southwest Forestry University, Kunming 650224, China
- <sup>5</sup> Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China
- \* Correspondence: zhangxc@ibcas.ac.cn (X.Z.); yhy@ibcas.ac.cn (Y.Y.)

**Abstract:** Although extant lycophytes represent the most ancient surviving lineage of early vascular plants, their plastomic diversity has long been neglected. The ancient evolutionary history and distinct genetic diversity patterns of the three lycophyte families, each with its own characteristics, provide an ideal opportunity to investigate the interfamilial relationships of lycophytes and their associated patterns of evolution. To compensate for the lack of data on Lycopodiaceae, we sequenced and assembled 14 new plastid genomes (plastomes). Combined with other lycophyte plastomes available online, we reconstructed the phylogenetic relationships of the extant lycophytes based on 93 plastomes. We analyzed, traced, and compared the plastomic diversity and divergence of the three lycophyte families (Isoëtaceae, Lycopodiaceae, and Selaginellaceae) in terms of plastomic diversity by comparing their plastome sizes, GC contents, substitution rates, structural rearrangements, divergence times, ancestral states, RNA editings, and gene losses. Comparative analysis of plastid phylogenomics and plastomic diversity of three lycophyte families will set a foundation for further studies in biology and evolution in lycophytes and therefore in vascular plants.

**Keywords:** lycophytes; plastid genome; comparative genomics; clubmoss; plastomic diversity



**Citation:** Chen, S.; Wang, T.; Shu, J.; Xiang, Q.; Yang, T.; Zhang, X.; Yan, Y. Plastid Phylogenomics and Plastomic Diversity of the Extant Lycophytes. *Genes* **2022**, *13*, 1280. <https://doi.org/10.3390/genes13071280>

Academic Editor: Bin Yu

Received: 24 June 2022

Accepted: 13 July 2022

Published: 19 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Vascular plants evolved from ancestral vascular-free land plants more than 400 million years ago [1,2], which is a major step during the colonization of land by plants [3,4]. As a sister group of all other extant terrestrial plants and the living clade most similar to early vascular plants of the fossil record, lycophytes retain many typical features of early vascular plants [5], and hence are suitable for investigating the early evolution of vascular plants. Although they dominated the earth's flora through the Devonian and Carboniferous periods [6,7], only three distinctive lineages (clubmosses, quillworts, and spikemosses) are extant, corresponding to three families: Lycopodiaceae, Isoëtaceae, and Selaginellaceae [8]. There are perhaps 700 species in Selaginellaceae within a single genus, about 388 species in Lycopodiaceae in 16 genera, and approximately 250 species in the single monogeneric family Isoëtaceae [8]. Their early divergence and evolutionary innovations enable their broad global distribution and diverse ecological niches, as well as abundant diversity at the genetic level [9].

Plastids are vital organelles for photosynthesis, amino acid and lipid synthesis, and many other biochemical pathways in plants. They provide energy to green plants and

play a vital role in the growth and development of plants. They are usually uniparentally inherited, conserved in structure, and easy to be sequenced for having a high copy number, thus being an ideal choice for plant phylogenetics and evolutionary inquiry [10]. The advent of next-generation sequencing techniques and open data sources makes it much easier for researchers to sequence, download, and analyze complete plastid genomes (plastomes), and with more and more plastome data available in databases, universal features of the plant plastome have been gradually revealed. Plastomes of most organisms have a quadripartite circular structure, ranging from 115 to 165 Kb in size [11,12]. In general, a typical quadripartite plastome consists of two inverted repeat (IR) sequences: a large single copy (LSC) and a small single copy (SSC). However, some exceptions exist. Loss of one of the IR regions is not rare in some clades of seed plants, such as Pinaceae [13], cupressophytes [14], Passiflora [15], and some clades of Fabaceae [16]. In lycophytes, a pair of direct repeats (DR) was found in most species of Selaginellaceae [17], while Lycopodiaceae and Isoëtaceae share typical IR structures [18]. As for other characteristics of plastomes, such as GC content, organelle genomes are near universally rich in AT throughout the eukaryotic domain, with an average GC content of 45% [19]. The highest plastid GC content holding the record appears in *Selaginella remotifolia*, at 56.5% [17], and the GC content of nearly all Selaginellaceae plastomes is extraordinarily high. Interestingly, despite their common ancient origins, the plastomes of the other two lycophyte families are much lower in GC content. As reported in previous studies, the GC content of Selaginellaceae plastomes ranges from 50.7% in *Selaginella lyalliito* to 56.5% in *S. remotifolia* [17]. However, the GC contents of Lycopodiaceae and Isoëtaceae are relatively low when compared throughout all ferns and lycophytes [20]: the GC content of Lycopodiaceae ranges from 35% to 36%, that of Isoëtaceae is approximately 38% [18], and that of ferns ranges from 33% to 45% in plastids. The plastid structure and features of the different families of the lycophytes differentiated in distinctive directions, becoming an intriguing evolutionary puzzle that has attracted researchers for a long time [17,18,21–23].

The exponential increase in the speed and accuracy of plastid sequencing has contributed to our understanding of lycophytes. To date, many DNA fragments have been applied to research on lycophytes [24–29], and since the publication of the first lycophyte plastome of *Huperzia lucidula* (Michx.) Trevis. (Lycopodiaceae) in 2005 [30], remarkable achievements have been made in the molecular evolution and classification of lycophytes. From 14 plastids reported in 2018 [18] to 81 complete lycophyte plastomes in this article (14 newly sequenced), combined with published nuclear genomes of *Selaginella kraussiana* (Kunze) A. Braun [5], *S. moellendorffii* Hieron. [31], *S. tamariscina* (P. Beauv.) Spring [32], and *Isoetes taiwanensis* De Vol [33], numerous mysteries of lycophytes have gradually been uncovered in recent years.

However, although several studies on lycophyte plastomes have been reported to date, they have covered relatively few species and have focused on only limited genera or families, resulting in an insufficient understanding of the lycophytes as a whole [17,18,21,22,34–36]. Indeed, compared with those of other terrestrial plants, the plastomes of lycophytes are relatively scarce, particularly for Lycopodiaceae. To explore the evolutionary patterns and strategies of the entire lycophyte lineage and the differences among the three lycophyte families, more samples and more species must be sequenced and investigated.

In this study, 14 plastomes were newly sequenced and assembled to compensate for the lack of data on Lycopodiaceae, thus enriching the plastome data of the lycophytes. The new plastomes were combined with 67 previously published lycophyte plastomes and 12 plastomes of other lineages to perform a comprehensive plastomic analysis of the lycophytes. In total, we analyzed 93 plastomes and aimed to (i) reconstruct the lycophyte phylogeny based on plastid protein-coding sequences; (ii) present the global structural characteristics of plastids from the three lycophyte families (Lycopodiaceae, Isoëtaceae, and Selaginellaceae); and (iii) characterize the features, differences, and evolutionary trajectory of lycophyte plastomes by comparing plastomic diversity of the different families in terms of GC content, variation of gene number, divergence time, substitution rate, RNA editing,

## 2. Materials and Methods

### 2.1. Taxon Sampling, DNA Extraction, and Sequencing

Fourteen individuals from 12 species, *Lycopodiastrium casuarinoides* (collection No. YYH15707, from Nanling, China), *Diphasiastrum complanatum* (YYH15882, from Bijie, China), *Lycopodium japonicum* (YYH15880, from Bijie, China), *Dendrolycopodium obscurum* (YYH15871, from Changbai Mountain, China), *Dendrolycopodium verticale* (YYH15881, from Bijie, China), *Pseudolycopodiella caroliniana* (YYH15570, from Yangjiang, China), *Palhinhaea cernua* (YYH15722 from Nanling, China; css1 from Shenzhen, China; YYH15886, from Jieyang, China), *Huperzia crispata* (146, from Hongya, China), *Huperzia selago* (3239, from Eastern Alps, Austria), *Phlegmariurus henryi* (YYH15875, from Fangchenggang, China), *Phlegmariurus mingcheensi* (YYH15514, from Dawuling Natural Reserve, China), and *Phlegmariurus squarrosus* (YYH15879, from Jingxi, China), were used for genome skimming and complete plastome assembly (Table S1). All leaf samples were frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . Total DNA was extracted from the leaves using a modified CTAB method and sequenced using short reads produced by the NovaSeq 6000 platform ( $2 \times 150$  bp) by Novogene (Beijing, China).

### 2.2. Plastome Assembly, and Annotation

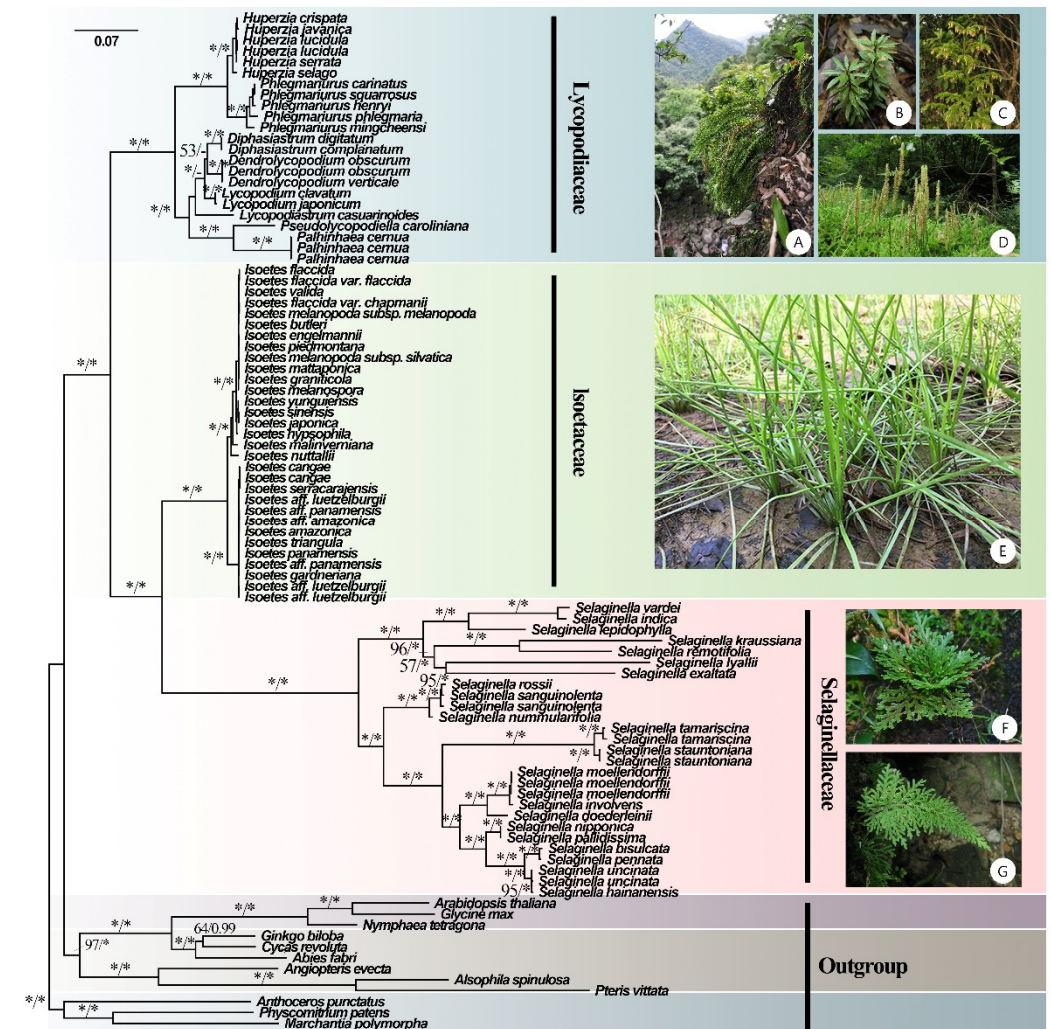
The raw reads were filtered using Trimmomatic 0.39 [37] by removing low-quality bases and sites, and the parameters were set as LEADING = 20, TRAILING = 20, SLIDING-WINDOW = 4:15, MINLEN = 36, and AVG QUAL = 20. The quality was assessed using FastQC 0.11.9 [38]. The filtered reads were assembled using the pipelines GetOrganelle 1.6.2 [39] with kmers: 21,45,65,85,105. Bandage 0.8.1 [40] was then used to visualize and complete the assembly. The annotation of plastomes was performed using PGA [41] and then visually inspected and edited by hand where necessary in Geneious v11.0.4 [42]. Published lycophyte plastomes (*Diphasiastrum digitatum*: MH549638, *D. obscurum*: MH549637, *Lycopodium clavatum*: MH549642, *H. lucidula*: AY660566 and MH549639, *Huperzia serrata*: KX426071, *H. javanica*: KY609860, *Phlegmariurus carinatus*: MN566837, *Phlegmariurus phlegmaria*: MT786212) are used as references in all assembly and annotation analysis.

### 2.3. Phylogenetic Analysis

This study included one genus of Isoëtaceae, one genus of Selaginellaceae, and eight genera representing all three subfamilies (Huperzioidae, Lycopodioidae, and Lycopodielloidae) of Lycopodiaceae. In addition to the 14 newly sequenced Lycopodiaceae plastomes, we downloaded the GenBank and Fasta format files for plastomes from 67 other lycophytes (31 Isoëtaceae, 27 Selaginellaceae, and 9 Lycopodiaceae) and 12 outgroup species (3 each of Bryophyta, Gymnospermae, Angiospermae, and Pteridophyta) from NCBI. GenBank accessions and more extensive data are listed in Supplementary Material Table S1. For plastomes with incomplete or no annotations, PGA was used for annotation, and Geneious 11.0.4 was used for visual inspection, with previously reported, closely related plastomes as references.

Phylogenetic analyses were conducted on 93 plastomes. First, all coding sequences (CDSs) were extracted in batches using the Python script `get_annotated_regions_from_gb.py` (<https://github.com/Kinggerm/PersonalUtilities/>, last accessed on 12 July 2022) [43]. A total of 84 CDSs with over 50% species coverage were selected. All sequences were aligned using Mafft 7.487 [44] and trimmed with Gblocks 0.91b [45,46], with default parameters to remove poorly aligned regions. Finally, we concatenated all CDSs into one matrix and partitioned them by codon position (1st, 2nd, and 3rd) in PhyloSuite 1.2.2 [47]. The best partitioning scheme and substitution model of the matrix was determined in PartitionFinder 2.0.0 [48] under the Bayesian information criterion (BIC), using the relaxed hierarchical clustering algorithm [49], and with the branch length linked. Based on the best partitioning scheme suggested by PartitionFinder, both Bayesian inference (BI) and maximum likelihood (ML) methods were applied in phylogenetic inference. A maximum likelihood (ML) tree was built using IQ-TREE 2.1.4 [50] with 1000 bootstrap replicates. Another phylogenetic

tree was constructed with MrBayes 3.2.7 [51]; two simultaneous runs were performed, each consisting of four chains, and the parameters were set to “lset nst = 6 rates = invgamma” for Bayesian inference (BI) analysis. In total, chains were run for 3,000,000 generations, with topologies sampled every 1000 generations. The first 25% of the sampled trees were discarded as burn-in. The convergence of runs was assumed when the average standard deviation of split frequencies dropped below 0.01. The best tree was edited using FigTree v1.4.4 (<http://tree.bio.ed.ac.uk>, last accessed on 12 July 2022), with ML bootstrap (BS) and Bayesian posterior probability (PP) support values shown for each node (Figure 1).



**Figure 1.** Phylogenetic tree inferred by the maximum likelihood (ML) and Bayesian methods based on 84 CDSs from 93 plastomes of 81 lycophytes and 12 outgroup species. ML bootstrap (BS) and Bayesian posterior probability (PP) support values are separated by ‘/’ and marked on each branch (\* indicates BS = 100 or PP = 1). (A) *Phlegmariurus petiolatus* [Photos: XL Zhou], (B) *Huperzia javanica* [Photos: SS Chen], (C) *Palhinhaea cernua* [Photos: YH Yan], (D) *L. japonicum* [Photos: YH Yan], (E) *Isoetes sinensis* [Photos: YF Gu], (F) *Selaginella tamariscina* [Photos: JP Shu], and (G) *Selaginella bififormis* [Photos: YH Yan].

#### 2.4. Comparative Plastomic Analysis

To focus on the evolutionary trajectories of the lycophytes, comparative plastomic analyses were performed on the lycophyte plastomes. The basic characteristics of lycophyte plastomes, such as genome length and GC content of overall, single copy, and long repeat regions, are listed in Supplementary Table S1. To compare the structure and gene contents of the plastomes, one of the repeat regions was excluded, and pairwise genome alignments

among the plastomes were performed separately for each family using progressive Mauve 2.4.0 [52].

The GC content along the complete sequence of each plastome was calculated with a sliding window size of 100 bp. Taking into account the particularity of repeat regions, separate sliding window detection was performed for repeat regions using the same sliding window size. Finally, eight widespread and representative species were selected for display because they could represent the common variation in each clade, and they are widely distributed and well-known species (*H. serrata*, *H. selago*, *P. phlegmaria*, *P. cernua*, *L. japonicum*, *Isoetes sinensis*, *Selaginella tamariscina*, and *S. lepidophylla*). Figures were prepared using the R language, with annotation information for each plastome obtained from Geneious 11.0.4 and depicted at the bottom of each scatter plot.

### 2.5. Divergence Time Estimation and Ancestral State Reconstruction

Before divergence time estimation and ancient trait reconstruction, we deleted all duplicate species, subspecies, varieties, forms, and species affinis, and kept only one plastome for each species. Finally, 71 plastomes from 93 plastomes were selected for analysis. With the same software and parameters described in Section 2.3, 84 CDSs were selected, aligned, and pruned into alignments. These alignments were concatenated and partitioned by codon position (1st, 2nd, and 3rd) in PhyloSuite 1.2.2 [47]. After the best partitioning scheme and substitution model were determined in PartitionFinder 2.0.0 [48], maximum likelihood (ML) inferences were performed using IQ-TREE 2.1.4 [50] with 1000 bootstrap replicates. This ML tree was used for divergence-time estimation and ancient trait reconstruction.

We estimated divergence times of the lycophytes with the MCMCTree program implemented in PAML 4.9 [53] based on the partitioned CDS dataset (partitioned by 1st, 2nd, and 3rd codon positions) and the species tree reconstructed above. MCMCTreeR was used to prepare the MCMCTree analyses and visualize the time-scaled trees [54]. Four fossils were used for calibration, following [2] and [36]. The divergence between Isoëtiales and Selaginellales was estimated to be 331–388 Ma, based on *Yuguangia ordinata* [55]. The node marking the divergence of heterosporous lycophytes (Isoëtaceae plus Selaginellaceae) and homosporous lycophytes (Lycopodiaceae) was constrained at 392–451 Ma based on *Leclercqia complexa* [56]; the crown node of Euphyllophyta was constrained at 386–451 Ma based on *Rellimia thomsonii* [57] and the Embryophyta crown node was estimated at 469–516 Ma based on *Tetrahedraletes cf. medinensis* [58], as proposed by Morris et al. [2]. Uniform distributions were applied to specify the prior distributions on node ages, allowing 0.001% probability of an age younger or older than the given minima and maxima. The time unit was set to 100 Ma. The clock and the model in the control file were set as independent rates and the GTR + G model (model = 7), respectively. The Markov chain Monte Carlo (MCMC) chains were run for 200 million generations, sampling every 100 generations, and the first 20 million generations were discarded as burn-in. Two independent runs were performed, and the stationary state and convergence of each run were assessed in Tracer 1.7.2 [59] to ensure sufficient and effective sample sizes for all parameters surpassing 200.

Ancestral trait reconstruction was performed to investigate how the gene number and GC content evolved across the lycophyte phylogeny. Ancestral GC contents and gene numbers were reconstructed using the function “ace” from the R package ape based on maximum likelihood (ML) methods on continuous characters [60] and mapped out through the corresponding branches of the divergence tree by the *ggtree* package in R [61].

### 2.6. Estimation of Substitution Rates

To compare substitution rates in the three families, the overall substitution rates (OSRs) of every ribosomal RNA (rRNA) gene and two noncoding region datasets (IR dataset and SC dataset) were estimated with BASEML in PAML 4.9 [53] using the GTR (REV) substitution model. For other genes, a concatenated matrix of all CDSs was used to estimate the synonymous substitution rate (dS) and nonsynonymous substitution rate (dN) with the codeml program in PAML 4.9 under the branch model (model = 2). Three branches

(corresponding to the three families) were regarded as having three separate  $\omega$  (dN/dS) values, but the same  $\omega$  value was shared within every family, and each “foreground” and “background” branch was specified. For the resulting OSR trees, dN trees, and dS trees, the branch lengths from the root to each tip were calculated using the R package *castor* 1.6.1 [62].

### 2.7. RNA Editing Analysis

A fraction of RNA editing would alter the sequences of start codons and stop codons, thus providing an opportunity for RNA editing inferences with a lack of mature mRNA or cDNA. To estimate the number of RNA-editing sites in each lycophyte plastome, the GFF3, GenBank, and Fasta files of all CDSs from each plastome were extracted separately, and the number and type of potential RNA-editing sites in the stop codons were then counted using the R package *ReFernment* [63]. *ReFernment* can automatically detect apparent nonsense mutations in plastid genomes to annotate hypothesized RNA-editing sites. Although it detects fewer RNA-editing sites than other similar tools, such as *PREPACT*, its results are more reliable and suitable for this research. Because some plastomes were published with the RNA editing corrected before uploading them to NCBI [64], the RNA-editing sites of a plastome that were obviously atypical compared with those of other species in the same family were treated as outliers and removed. For two of the outliers in *Isoëtaceae*, *Isoëtes hypsophila* and *Isoëtes japonica* [64], we contacted the authors and obtained the original plastome sequences.

## 3. Results

### 3.1. Features of Lycophyte Plastomes

Fourteen new plastomes from 12 species of *Lycopodiaceae* were sequenced, assembled, and annotated in this study. Detailed information on these plastomes is listed in Table S1, together with information on the other 79 downloaded plastomes. These newly sequenced plastomes exhibit a typical quadripartite structure: a large single copy region (LSC), a short single copy region (SSC), and two inverted repeats (IRs). The lengths of these complete plastomes range from 145,398 bp (*P. cernua*) to 168,075 bp (*Huperzia selago*), and their overall GC contents range from 32.9% (*L. casuarinoides*) to 36.3% (*Huperzia crispata*).

There are obvious differences in the plastomes of the three families. In terms of GC content, that of *Lycopodiaceae* was the lowest (32.9–36.4%), followed by *Isoëtaceae* (37.7–38.2%) and *Selaginellaceae* (50.2–56.5%). In terms of plastome size, those of *Selaginellaceae* are the shortest, from 110,164 bp (*Selaginella nipponica*) to 148,924 bp (*S. sanguinolenta*), followed by those of *Isoëtaceae*, from 142,880 bp (*Isoëtes* aff. *amazonica*) to 146,362 bp (*I. hypsophila*), and *Lycopodiaceae*, from 145,297 bp (*P. cernua*) to 168,075 bp (*Huperzia selago*). A reduction in genome size was associated with a reduction in gene number. *Selaginellaceae* plastomes have the smallest number of genes, with the fewest in *Selaginella pallidissima* (64 genes) and the most in *S. sanguinolenta* (113 genes), followed by *Isoëtaceae* (129 genes on average), and *Lycopodiaceae* (131 genes on average) (Table S1 and Figure S1). In general, *Selaginellaceae* plastomes had the lowest gene numbers, the shortest lengths, and the highest GC contents; *Lycopodiaceae* plastomes had the largest gene numbers, the greatest lengths, and the lowest GC contents; and the *Isoëtaceae* plastomes were intermediate.

Plastome rearrangement analysis was performed by family, and the results show that there has been little genome rearrangement in *Lycopodiaceae* and *Isoëtaceae* (Figure S2A,B), especially *Isoëtaceae*, which showed almost no structural rearrangement or extreme structural conservation (Figure S2A). However, significant structural rearrangements appeared in *Selaginellaceae* (Figure S2C). Unlike *Selaginellaceae*, the other two families show normal IR structures.

### 3.2. Phylogenetic Analyses

A robust phylogenetic tree of lycophytes was constructed based on 81 plastomes from lycophytes and 12 from other land plants (Figure 1). In our phylogenomic analyses,

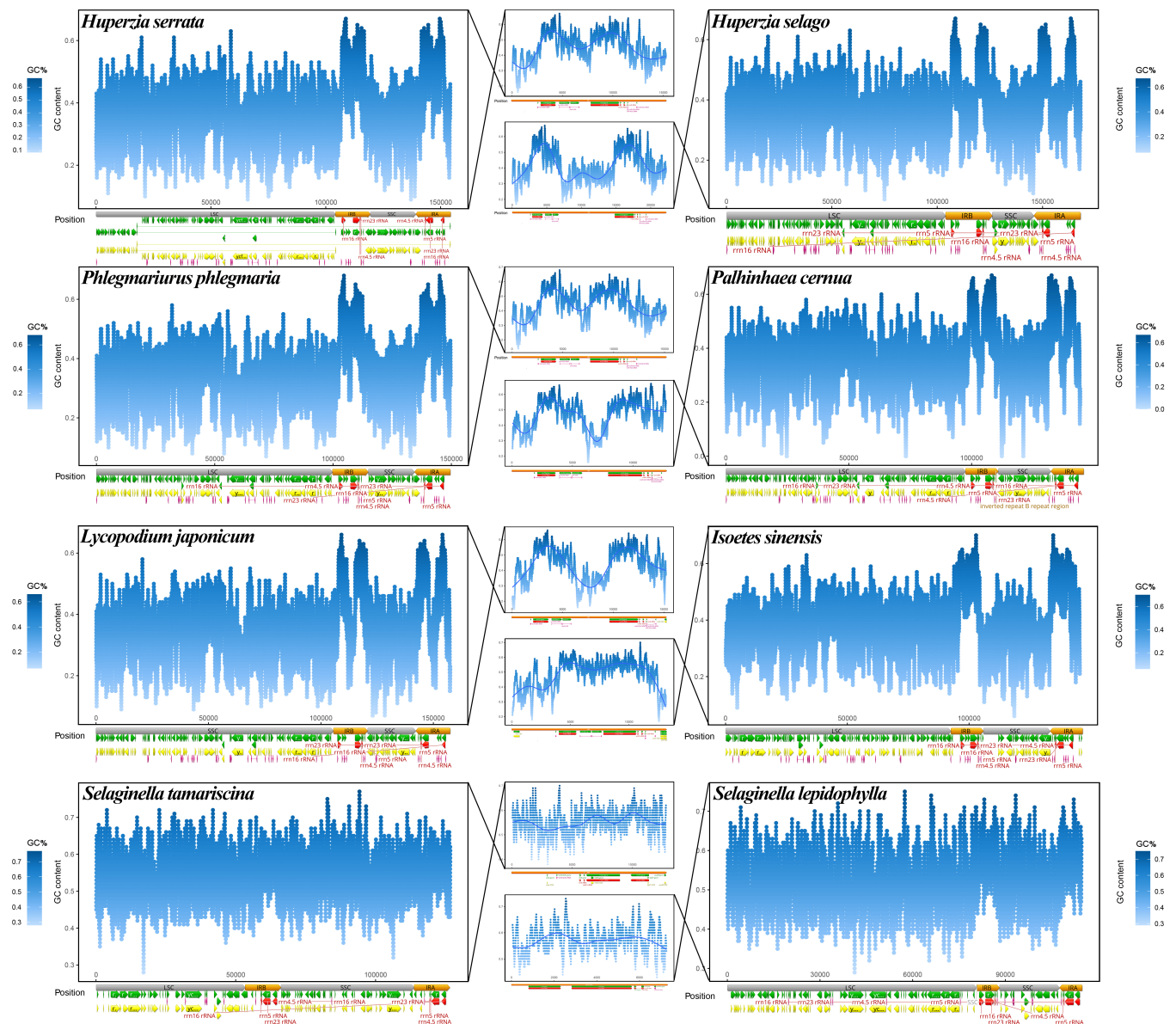
Lycophytes are sister to all other extant vascular plants and can be divided into three clades that correspond to three families (Lycopodiaceae, Isoëtaceae, and Selaginellaceae). BI and ML analyses recovered highly supported tree topologies, with the shortest branch lengths in Isoëtaceae, intermediate in Lycopodiaceae, and longest in Selaginellaceae (Figure 1). The phylogeny reconstructed the topology of eight genera of Lycopodiaceae: *Huperzia*, *Phlegmariurus*, *Dendrolycopodium*, *Diphasiastrum*, *Palhinhaea*, *Pseudolycopodiella*, *Lycopodias-trum*, and *Lycopodium*. However, the topology of the genus *Dendrolycopodium* is slightly discordant between the ML tree and the BI tree. In the BI tree, the genus *Diphasiastrum* sisters to *Dendrolycopodium*, while *Diphasiastrum* sisters to *Lycopodium* instead in the ML tree, both with low support values (0.64 for BI and 53 for ML) (Figure 1). As for Isoëtaceae and Selaginellaceae, their topology was divided into two lineages.

### 3.3. GC Content Distribution among Plastomes

To investigate the specific location of GC differences on different plastomes, we performed a sliding-window analysis of GC content in each plastome and detected GC peaks in rRNA genes and their adjacent regions. The GC distributions along eight representative plastomes are shown in Figure 2. In Isoëtaceae and Lycopodiaceae, the GC content was highest in the repeat regions, with two GC peaks in the IR regions, as shown in the middle panels in Figure 2. These peaks clearly corresponded to regions containing rRNA genes. The GC content of the repeat regions was slightly higher in Selaginellaceae than in the six Isoëtaceae and Lycopodiaceae plastomes in Figure 2 (Table S1, Figure S2), but in Selaginellaceae, the two GC peaks were not as prominent as in Isoëtaceae and Lycopodiaceae, due to the smaller GC difference between repeat regions and single copy regions. *Selaginella tamariscina*, with the DR structure, and *S. lepidophylla*, with the IR structure, shared a similar pattern in which the GC contents of the SC region and repeat region were nearly identical, regardless of the order of the two repeats. In conclusion, as well as an increase in the GC content of the repeat regions, Selaginellaceae showed an increase in the GC content of the SC region to a level similar to, though still not as high as that of the two repeat regions (Table S1). In contrast, the GC content of the IR regions was much higher than that of the SC region in the two other families of lycophytes.

### 3.4. Base Substitution Rate Estimation

As shown in Figure 3, the mutation rates of each family showed entirely different patterns. Within each family, the nonsynonymous substitution rate (dN) is always lower than the synonymous substitution rate (dS), and the coding regions always mutate more slowly than the noncoding regions. Among the three families, whether for dN or dS, in noncoding regions (OSR) or coding regions, and in repeat or single-copy regions, the substitution rate is generally much higher in Selaginellaceae than in Lycopodiaceae and Isoëtaceae. In addition, the substitution rates in the SC and IR regions increased synchronously, with substitution rates always lower in the two repeat regions than in the SC region. rRNA genes have extremely low substitution rates and extremely high GC contents, and peaks in GC content in the IR regions of plastomes are clearly caused by the presence of these genes (Figure 2). These results indicate low, intermediate, and high rates of plastomic change in the plastids of Isoëtaceae, Lycopodiaceae, and Selaginellaceae, respectively.

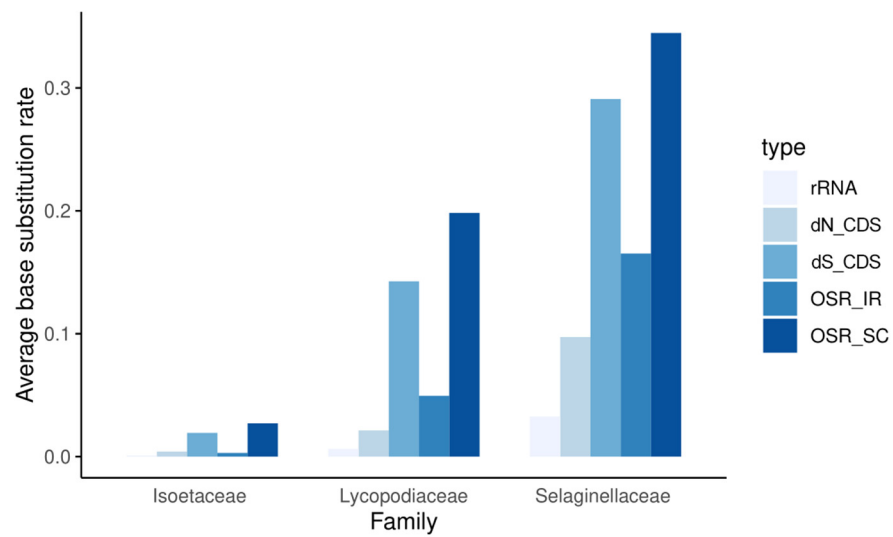


**Figure 2.** Sliding window analysis of GC content of 8 representative lycophte plastomes. Each large figure represents the GC content distribution along a plastome. The yellow, green, red, and orange blocks at the bottom of the blue plots represent genes, CDSs, tRNA, and rRNA genes, and the two repeat regions, respectively. The middle panels represent the GC content of one of the repeat regions, connected to the corresponding species by black crossed lines. All sliding window sizes were 100 bp.

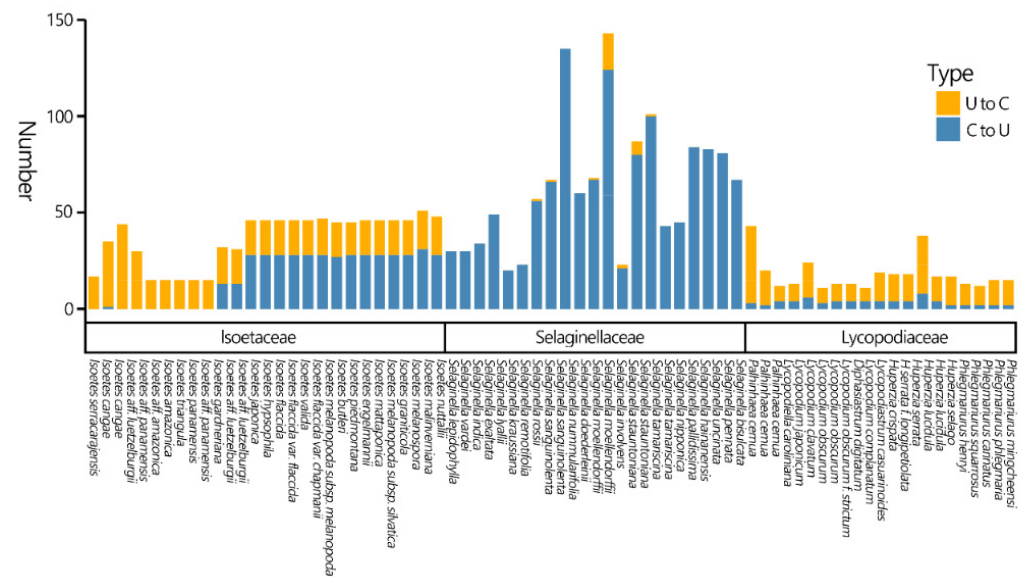
### 3.5. RNA Editing Analysis

In this study, we predicted only the numbers and types of RNA editing in CDSs. Among the three families, Selaginellaceae showed higher levels of C-to-U RNA editing (Figure 4). Plastomes of Lycopodiaceae contain both C-to-U and U-to-C RNA-editing sites, and the latter are far more common. Two completely different patterns appear in Isoëtaceae: some exhibit almost no C-to-U RNA editing, while others show both types of editing, with C-to-U predominating. The Selaginellaceae is estimated to have the most extensive C-to-U RNA editing numbers. Followed by Isoëtaceae splitting into two clades, both have similar U-to-C editing numbers, but one clade exhibits elevated C-to-U editing sites. Lycopodiaceae has the lowest number of RNA editing sites.





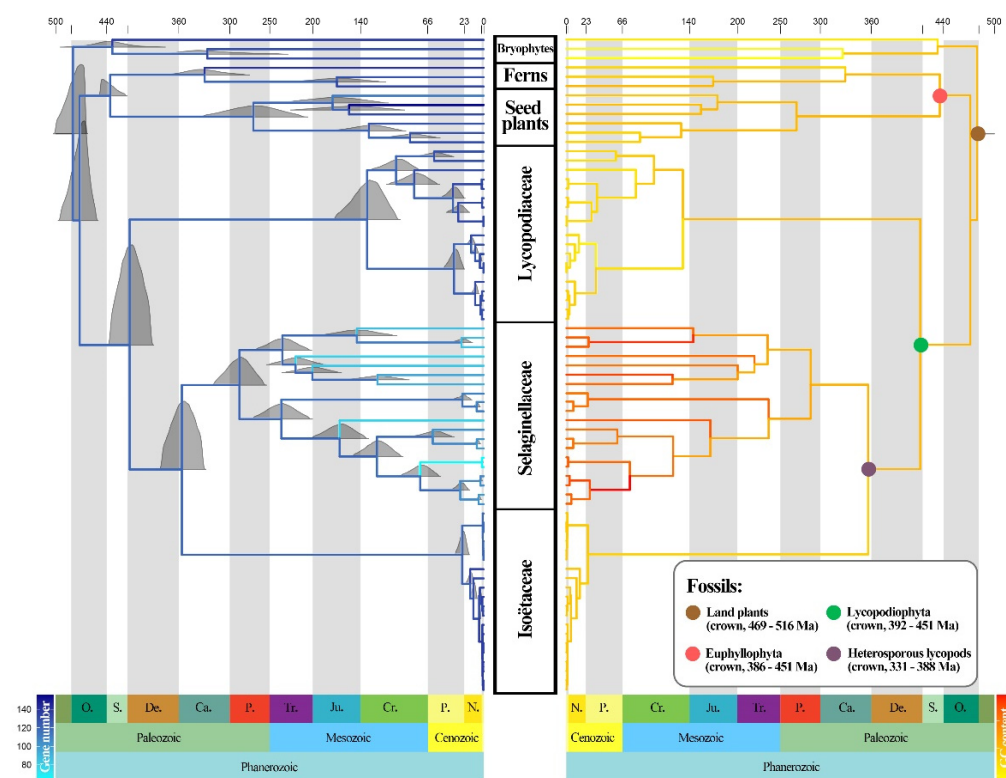
**Figure 3.** Overall substitution rates (OSRs) of noncoding sequences in repeat regions, SC regions, and rRNA regions; synonymous substitution rates (dS), nonsynonymous substitution rates (dN) of all coding sequences (CDS).



**Figure 4.** The numbers and types of predicted RNA-editing sites in the CDSs of lycoplyte plastomes.

### 3.6. Divergence Time Inference and Ancestral State Reconstruction

The divergence time inference and the ancestral state reconstruction of GC content and gene number in the lycoplytes are shown in Figure 5. Based on the four calibration points derived from fossils, our analysis estimated that the stem age of the lycoplytes was dated to 417 Ma in the Lower Devonian, with a 95% highest posterior density interval (HPD) of 394–443 Ma. Selaginellaceae and Isoëtaceae separated at around 355 Ma in the Carboniferous (95% HPD: 333–382). These two families would have started their diversification at approximately 287 Ma (95% HPD: 259–317) for Selaginellaceae and, more recently, during the early Paleogene at around 25 Ma (95% HPD: 19–34) for Isoëtaceae. Lycopodiaceae started to diversify during the Cretaceous at 137 Ma (95% HPD: 105–177).



**Figure 5.** Ancestral state reconstruction of gene numbers (left) and GC content (right) along the divergence time tree of the lycophytes inferred from MCMCTree. The full posterior distributions were displayed on nodes (left), and the positions of fossil node calibrations were marked by colored circles (right).

Ancestral state reconstruction showed that the three families experienced quite dissimilar evolutionary patterns in GC content and gene numbers. With the diversification of Lycopodiaceae, the GC content decreased over time, especially in *Phlegmariurus*. In contrast, the GC content of Isoëtaceae did not change greatly during diversification. Selaginellaceae showed a marked increase in GC content over time, accompanied by gene loss, whereas the number of genes in the other two families did not show a clear increasing or decreasing trend.

## 4. Discussion

### 4.1. Plastomic Diversity of Lycophytes

**Plastome Structure Divergence.** The plastomic diversity of lycophytes is reflected in two distinctive types of quadripartite structures: one is the IR structure shared by all Lycopodiaceae and Isoëtaceae species, and the other is the DR structure in most Selaginellaceae species. In-depth research on the plastome structure of Selaginellaceae shows that they include not only those with DR structures but also those in which both DR and IR structures coexist [17,21]. The DR and IR structures of the latter are reported to be dynamic, with recombination activities between repeats in plastomes giving rise to subgenomic molecules [17]. Ancestral state reconstruction analysis indicated that the DR structure was the ancestral state, whereas IR was derived in Selaginellaceae [17].

In terms of structure variation, through rearrangement analysis of each family, we found that the plastid structure of Isoëtaceae is more conservative than that of Lycopodiaceae, and much more conservative than Selaginellaceae, which is reflected in Figure 2: whole genome alignment of Isoëtaceae identified 2 locally collinear blocks (LCBs), while Lycopodiaceae identified 5, and Selaginellaceae identified 15. In previous studies, rearrangement analysis of genes also supported this conclusion [17,32]. The structural instability of Selaginellaceae plastomes is strongly associated with the DR structure and possible subge-

nomic molecules [17]. Most species of Selaginellaceae, or ancestor of Selaginellaceae, do not have an IR structure; however, it is generally believed that plastomes with IR structures are more stable, and plastid rearrangements are more frequent when a large inverted repeat sequence has been lost [65].

**Base Composition and Substitution Rate Divergence.** In terms of GC content diversity, the plastids of the lycophytes almost evolved in two opposite directions. Plastid genomes of almost all Selaginellaceae possess extraordinarily high GC contents, and the GC content of their another organelle, the mitochondrion, is also GC-rich, with the highest documented mitochondrial GC content (68%) appearing in *S. moellendorffii* [66], but there is nothing unusual about the GC content of nuclear genomes in Selaginellaceae, all of which are approximately 50% and unremarkable compared with other nuclear genomes of other green plants [5,31,32]. In contrast, the GC content of Lycopodiaceae and Isoëtaceae are relatively low throughout all ferns and lycophytes [20].

The different base substitution rates of lycophytes might reflect their plastomic diversity and are related to GC content. Overall, the base substitution rates of Isoëtaceae are slower than those of Lycopodiaceae, and those of Lycopodiaceae are slower than those of Selaginellaceae. As observed in *Arabidopsis thaliana*, mutations are neither random nor evenly distributed in the genome; mutation rates are much slower in functionally constrained regions of the genome [67], which may explain the especially low base substitution rates of rRNA genes (Figure 2), as they are functionally constrained. The nuclear genome is GC-rich in rRNA loci [68], and regions near the rRNA genes also have high GC content. High GC content may reflect the nature of the rRNA gene itself, as even in prokaryotes with generally fast mutation rates, their rRNAs are extremely conserved and GC rich [69]. Therefore, the repeat region, composed of rRNA genes and their nearby areas, exhibits the highest GC content and the lowest mutation rate in the plastome because of its extremely high conservation. Since rRNA genes are generally located in the repeat regions, two repeat regions become the most GC-rich among the four sections of every quadripartite plastome, regardless of DR or IR structure. However, the GC contents of both repeat regions and single-copy regions of Selaginellaceae are almost as high as those of repeat regions, which leads to the high overall GC content of Selaginellaceae directly.

**RNA Editing Divergence.** Phylogenetic distribution of C-to-U and U-to-C RNA editing has already been found among land plants, which is related to the origins and evolutionary trajectories of two types of RNA editing [70]. In general, C-to-U editing exists in the mitochondria and plastids of all major land plant groups, while not in closely allied green algal lineages or in marchantiid liverworts, and U-to-C editing appears only in ferns, lycophytes, and hornworts [70]. Here, we found several factors influencing the RNA editing of the lycophyte. As suggested in Figure 4, different taxonomic groups have different RNA editing types and numbers. Selaginellaceae plastomes show high levels of C-to-U RNA editing, consistent with the high occurrence of C-to-U RNA editing reported in *Selaginella* plastomes [71,72]. By contrast, we found that Lycopodiaceae mainly shows U-to-C RNA editing, and two completely different clades are present in Isoëtaceae plastomes: one shows almost all C-to-U editing, and the other shows two types coexist.

We also found an effect of mutation on RNA editing. In terms of the number of RNA-editing sites, the number of sites in Lycopodiaceae was less than that in Isoëtaceae (Figure 3). This may be related to mutations at the DNA level, which are negatively correlated with the number of RNA-editing sites. The decrease in the general mutation rate may result in an increase in the number of editing sites that are retained because edited sites could have been lost due to mutations over time, as hypothesized by Tillich et al. based on evidence from seed plants [73]. Therefore, Isoëtaceae plastomes with slower base substitution rates are more likely to retain more RNA-editing sites than Lycopodiaceae plastomes (Figures 3 and 4). However, given the particularity of the high occurrence of RNA editing in both *Selaginella* plastid and mitochondrial genomes [66,71,72], this model cannot account for the RNA editing of Selaginellaceae. Instead, some site recognition factors for editing events in the *Selaginella* genome may be associated with the abundance

of organelle RNA editing [74]. For example, an expanded DYW-type PPR protein family (one of the key players of RNA editing [75]) is found in the *S. moellendorffii* genome [31].

#### 4.2. Phylogenetic Relationships and Divergence Times

Here, we present a robust lycophyte phylogeny yielded from a phylogenomic analysis based on 93 plastomes (Figure 1). The topology of the ML tree was highly consistent with that of the BI tree, except for disagreement on the phylogenetic position of *Dendrolycopodium* and *Lycopodium* (Figure 1). Before this study, only a few plastid markers (*atpA*, *psbA-trnH*, *rbcL*, *rps4* and *rps4-trnS*, *trnL* and *trnL-F*) or a few species were analyzed in phylogeny studies of Lycopodiaceae [8,18,29]. Further phylogenetic work under global and broad sampling and sequencing of Lycopodiaceae and the lycophyte is needed.

The divergence times of the major lineages of the lycophytes were near the boundaries of the geological periods (Figure 5). The first lycophytes can appear at the end of the Silurian and the beginning of the Devonian, as shown in our results (Figure 5) and supported by fossil evidence (Baragwanathia and Asteroxylon) [76,77] and many other divergence time analyses [17,36]. The divergence of the living homosporous and heterosporous lycophytes (Selaginellaceae and Isoëtaceae) was estimated to be during the early Carboniferous and late Devonian (355 Ma, 95% HPD: 333–382), similar to the results of previous molecular phylogenetic works [17,36,78]. Since the Devonian–Carboniferous boundary is associated with the second of the “big five” mass extinction events of Earth’s history, this divergence can be speculated to be infected by the joint effect of geological and climate changes. Among the three families of the lycophytes, Selaginellaceae was the first to diversify among the lycophytes during Permian and Carboniferous (287 Ma, 95% HPD: 259–317). This was followed by divergences among subfamilies of the involved Lycopodiaceae, which were dated to be during the Cretaceous and Jurassic (137 Ma, 95% HPD: 105–177). The crown age of Isoëtaceae is estimated to be 25 Ma (95% HPD: 19–34), slightly older than the age estimated by Pereira et al. [36]. The diversification of Isoëtaceae appeared recently, with a burst of species diversified in a short period of time, implying a recent evolutionary radiation, and so on in *Phlegmariurus* and *Huperzia* in Lycopodiaceae (Figure 5). This age estimate is consistent with the relatively recent radiation reported in previous phylogeny studies of Isoëtaceae [35,36] and *Phlegmariurus* [27,79].

#### 5. Conclusions

Lycophytes provide a unique opportunity to study the evolution of vascular plants. However, at present, the overall genetic diversity of all lycophytes remains poorly understood. We sequenced and assembled 14 new plastomes of Lycopodiaceae species, and combined the new sequencing data with published plastomes to construct a relatively complete phylogenetic relationship of 81 lycophytes and 12 outgroups. This study explored the plastid phylogenomics, plastomic structural diversity, and compositional variation of lycophytes by assessing differences in substitution rate, GC content, genome rearrangement, RNA editing, and gene loss based on the plastomic datasets. Among the three families, Selaginellaceae diversified first and exhibited the most variable plastome structure and high GC content, together with reduced gene numbers and RNA editing sites in their plastomes; Isoëtaceae experienced recent radiation, with the lowest base substitution rate but a considerable number of RNA editing sites retained. Lycopodiaceae is also conserved in structure and exhibits the lowest GC content but the largest number of genes. Determining the plastid phylogenomics and plastomic diversity of the lycophyte will not only deepen our understanding of the evolution of these early vascular plants but will also benefit research on all terrestrial plants. However, the evolution of lycophytes cannot be resolved using plastomes alone, and more mechanisms related to the evolution of lycophytes, such as inheritance, codon usage, and adaptive evolution, remain to be studied with more approaches and samples.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13071280/s1>, Table S1: Sample information and basic characteristics of 81 Lycophyte and 12 outgroup plastomes; Figure S1: Boxplot showing variation of GC content, genome size, and gene number of three families of the lycophyte; Figure S2: MAUVE alignment of the plastomes of three families; Files S1. Best partition schemes of PartitionFinder2 used in the phylogenetic analyses (IQtree and MrBayes).

**Author Contributions:** Conceptualization, Q.X. and Y.Y.; Formal analysis, S.C. and T.W.; Funding acquisition, X.Z. and Y.Y.; Investigation, S.C. and J.S.; Methodology, T.Y.; Supervision, X.Z. and Y.Y.; Writing—original draft, S.C.; Writing—review & editing, T.W., J.S., T.Y. and Y.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the China National Orchid Conservation Center and the Orchid Conservation and Research Center of Shenzhen (CNOCC), State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, the Chinese Academy of Sciences (LSEB, IBCAS), and University of Chinese Academy of Sciences (UCAS). This work was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA19050404), National Natural Science Foundation of China (31872651), and Basic and Applied Basic Research Foundation of Guangdong Province (2021A1515010920).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are openly available in the National Center for Biotechnology Information under the accession numbers presented in Table S1.

**Acknowledgments:** We are grateful to Yufeng Gu for sharing plastid genomes and providing specimens.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kenrick, P.; Crane, P.R. The origin and early evolution of plants on land. *Nature* **1997**, *389*, 33–39. [[CrossRef](#)]
2. Morris, J.L.; Puttick, M.N.; Clark, J.W.; Edwards, D.; Kenrick, P.; Pressel, S.; Wellman, C.H.; Yang, Z.; Schneider, H.; Donoghue, P.C.J. The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E2274–E2283. [[CrossRef](#)] [[PubMed](#)]
3. Bennici, A. Unresolved problems on the origin and early evolution of land plants. *Riv. Biol.* **2007**, *100*, 55–67. [[PubMed](#)]
4. Bennici, A. Origin and early evolution of land plants. *Commun. Integr. Biol.* **2008**, *1*, 212–218. [[CrossRef](#)] [[PubMed](#)]
5. Ge, Y.; Liu, J.; Zeng, M.; He, J.; Qin, P.; Huang, H.; Xu, L. Identification of WOX family genes in *Selaginella kraussiana* for studies on stem cells and regeneration in lycophytes. *Front. Plant Sci.* **2016**, *7*, 93. [[CrossRef](#)]
6. Kenrick, P.; Strullu-Derrien, C. The origin and early evolution of roots. *Plant Physiol.* **2014**, *166*, 570–580. [[CrossRef](#)]
7. White, M.E. *The Greening of Gondwana*; Reed: Sydney, Australia, 1986.
8. Schuettelpelz, E.; Schneider, H.; Smith, A.R.; Hovenkamp, P.; Prado, J.; Rouhan, G.; Salino, A.; Sundue, M.; Almeida, T.E.; Parris, B.; et al. A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* **2016**, *54*, 563–603. [[CrossRef](#)]
9. Spencer, V.; Nemeček, Z.; Harrison, C.J. What can lycophytes teach us about plant evolution and development? *Modern perspectives on an ancient lineage. Evol. Dev.* **2020**, *23*, 174–196.
10. Wolfe, A.D.; Randle, C.P. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: Implications for plant molecular systematics. *Syst. Bot.* **2004**, *29*, 1011–1020. [[CrossRef](#)]
11. Wolfe, K.H.; Li, W.H.; Sharp, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9054–9058. [[CrossRef](#)]
12. Jansen, R.K.; Raubeson, L.A.; Boore, J.L.; Depamphilis, C.W.; Chumley, T.W.; Haberle, R.C.; Wyman, S.K.; Alverson, A.J.; Peery, R.; Herman, S.J.; et al. Methods for obtaining and analyzing whole chloroplast genome sequences. *Methods Enzymol.* **2005**, *395*, 348–384. [[PubMed](#)]
13. Wu, C.S.; Wang, Y.N.; Hsu, C.Y.; Lin, C.P.; Chaw, S.M. Loss of different inverted repeat copies from the chloroplast genomes of pinaceae and cupressophytes and influence of heterotachy on the evaluation of gymnosperm phylogeny. *Genome Biol. Evol.* **2011**, *3*, 1284–1295. [[CrossRef](#)]
14. Wu, C.S.; Chaw, S.M. Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): Evolution towards shorter intergenic spacers. *Plant Biotechnol. J.* **2014**, *12*, 344–353. [[CrossRef](#)] [[PubMed](#)]
15. dos Santos, L.A.C.; Da Costa, Z.P.; Callot, C.; Cauet, S.; Zucchi, M.I.; Bergès, H.; Berg, C.V.D.; Vieira, M.L.C. A repertoire of rearrangements and the loss of an inverted repeat region in *Passiflora* chloroplast genomes. *Genome Biol. Evol.* **2020**, *12*, 1841–1857. [[CrossRef](#)] [[PubMed](#)]
16. Cronk, Q.; Ojeda, I.; Pennington, R.T. Legume comparative genomics: Progress in phylogenetics and phylogenomics. *Curr. Opin. Plant Biol.* **2006**, *9*, 99–103. [[CrossRef](#)]

17. Zhang, H.R.; Xiang, Q.P.; Zhang, X.C. The unique evolutionary trajectory and dynamic conformations of DR and IR/DR-coexisting plastomes of the early vascular plant *Selaginellaceae* (Lycophyte). *Genome Biol. Evol.* **2019**, *11*, 1258–1274. [[CrossRef](#)]
18. Mower, J.P.; Ma, P.; Grewe, F.; Taylor, A.; Michael, T.P.; VanBuren, R.; Qiu, Y. Lycophyte plastid genomics: Extreme variation in GC, gene and intron content and multiple inversions between a direct and inverted orientation of the rRNA repeat. *New Phytol.* **2019**, *222*, 1061–1075. [[CrossRef](#)]
19. Smith, D.R. Updating our view of organelle genome nucleotide landscape. *Front. Genet.* **2012**, *3*, 175. [[CrossRef](#)]
20. Du, X.; Lu, J.; Li, D. Advances in the evolution of plastid genome structure in lycophytes and ferns. *Biodivers. Sci.* **2019**, *27*, 1172–1183.
21. Kang, J.S.; Zhang, H.R.; Wang, Y.R.; Liang, S.Q.; Mao, Z.Y.; Zhang, X.C.; Xiang, Q. Distinctive evolutionary pattern of organelle genomes linked to the nuclear genome in Selaginellaceae. *Plant J.* **2020**, *104*, 1657–1672. [[CrossRef](#)]
22. Shim, H.; Lee, H.J.; Lee, J.; Lee, H.O.; Kim, J.H.; Yang, T.J.; Kim, N.-S. Plastid genomes of the early vascular plant genus *Selaginella* have unusual direct repeat structures and drastically reduced gene numbers. *Int. J. Mol. Sci.* **2021**, *22*, 641. [[CrossRef](#)]
23. Smith, D.R. Unparalleled GC content in the plastid DNA of *Selaginella*. *Plant Mol. Biol.* **2009**, *71*, 627–639. [[CrossRef](#)] [[PubMed](#)]
24. Small, R.L.; Lickey, E.B.; Shaw, J.; Hauk, W.D. Amplification of noncoding chloroplast DNA for phylogenetic studies in lycophytes and monilophytes with a comparative example of relative phylogenetic utility from Ophioglossaceae. *Mol. Phylogenet. Evol.* **2005**, *36*, 509–522. [[CrossRef](#)]
25. Zhou, X.M.; Rothfels, C.J.; Zhang, L.; He, Z.R.; Le Péchon, T.; He, H.; Lu, N.T.; Knapp, R.; Lorence, D.; He, X.-J.; et al. A large-scale phylogeny of the lycophyte genus *Selaginella* (*Selaginellaceae: Lycopodiopsida*) based on plastid and nuclear loci. *Cladistics* **2016**, *32*, 360–389. [[CrossRef](#)] [[PubMed](#)]
26. Pereira, J.B.S.; Labiak, P.H.; Stützel, T.; Schulz, C. Origin and biogeography of the ancient genus *Isoëtes* with focus on the Neotropics. *Bot. J. Linn. Soc.* **2017**, *185*, 253–271. [[CrossRef](#)]
27. Bauret, L.; Field, A.R.; Gaudeul, M.; Selosse, M.A.; Rouhan, G. First insights on the biogeographical history of *Phlegmariurus* (*Lycopodiaceae*), with a focus on Madagascar. *Mol. Phylogenet. Evol.* **2018**, *127*, 488–501. [[CrossRef](#)]
28. Testo, W.L.; Sessa, E.; Barrington, D.S. The rise of the Andes promoted rapid diversification in Neotropical *Phlegmariurus* (*Lycopodiaceae*). *New Phytol.* **2019**, *222*, 604–613. [[CrossRef](#)] [[PubMed](#)]
29. Chen, D.; Zhou, X.; Rothfels, C.J.; Shepherd, L.D.; Knapp, R.; Zhang, L.; Lu, N.T.; Fan, X.; Wan, X.; Gao, X.; et al. A global phylogeny of *Lycopodiaceae* (Lycopodiales; lycophytes) with the description of a new genus, *Brownseya*, from Oceania. *Taxon* **2022**, *71*, 25–51. [[CrossRef](#)]
30. Wolf, P.G.; Karol, K.G.; Mandoli, D.F.; Kuehl, J.; Arumuganathan, K.; Ellis, M.W.; Mishler, B.D.; Kelch, D.G.; Olmstead, R.G.; Boore, J.L. The first complete chloroplast genome sequence of a lycophyte, *Huperzia lucidula* (*Lycopodiaceae*). *Gene* **2005**, *350*, 117–128. [[CrossRef](#)]
31. Banks, J.A.; Nishiyama, T.; Hasebe, M.; Bowman, J.L.; Gribskov, M.; DePamphilis, C.; Albert, V.A.; Aono, N.; Aoyama, T.; Ambrose, B.A.; et al. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **2011**, *332*, 960–963. [[CrossRef](#)] [[PubMed](#)]
32. Xu, Z.; Xin, T.; Bartels, D.; Li, Y.; Gu, W.; Yao, H.; Liu, S.; Yu, H.; Pu, X.; Zhou, J.; et al. Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. *Mol. Plant* **2018**, *11*, 983–994. [[CrossRef](#)]
33. Wickell, D.; Kuo, L.-Y.; Yang, H.-P.; Ashok, A.D.; Irisarri, I.; Dadras, A.; de Vries, S.; de Vries, J.; Huang, Y.-M.; Li, Z.; et al. Underwater CAM photosynthesis elucidated by *Isoëtes* genome. *Nat. Commun.* **2021**, *12*, 6348. [[CrossRef](#)]
34. Nunes, G.; Oliveira, R.R.M.; Guimarães, J.T.F.; Giulietti, A.M.; Caldeira, C.; Vasconcelos, S.; Pires, E.; Dias, M.; Watanabe, M.; Pereira, J.; et al. Quillworts from the Amazon: A multidisciplinary populational study on *Isoëtes serracarajensis* and *Isoëtes cangae*. *PLoS ONE* **2018**, *13*, e0201417. [[CrossRef](#)]
35. Schafran, P.W.; Zimmer, E.A.; Taylor, W.C.; Musselman, L.J. A whole chloroplast genome phylogeny of diploid species of *Isoëtes* (*Isoëtaceae*, *Lycopodiophyta*) in the southeastern United States. *Castanea* **2018**, *83*, 224–235. [[CrossRef](#)]
36. Pereira, J.B.S.; Giulietti, A.M.; Prado, J.; Vasconcelos, S.; Watanabe, M.T.C.; Pinangé, D.S.B.; Oliveira, R.R.M.; Pires, E.D.; Caldeira, C.F.; Oliveira, G. Plastome-based phylogenomics elucidate relationships in rare *Isoëtes* species groups from the Neotropics. *Mol. Phylogenet. Evol.* **2021**, *161*, 107177. [[CrossRef](#)]
37. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
38. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 21 April 2020).
39. Jin, J.J.; Yu, W.B.; Yang, J.B.; Song, Y.; Depamphilis, C.W.; Yi, T.S.; Li, D.-Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [[CrossRef](#)] [[PubMed](#)]
40. Wick, R.R.; Schultz, M.B.; Zobel, J.; Holt, K.E. Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics* **2015**, *31*, 3350–3352. [[CrossRef](#)]
41. Qu, X.J.; Moore, M.J.; Li, D.Z.; Yi, T.S. PGA: A software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* **2019**, *15*, 50. [[CrossRef](#)]

42. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)] [[PubMed](#)]
43. Zhang, R.; Wang, Y.H.; Jin, J.J.; Stull, G.W.; Bruneau, A.; Cardoso, D.; De Queiroz, L.P.; Moore, M.J.; Zhang, S.-D.; Chen, S.-Y.; et al. Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst. Biol.* **2020**, *69*, 613–622. [[CrossRef](#)] [[PubMed](#)]
44. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
45. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **2000**, *17*, 540–552. [[CrossRef](#)] [[PubMed](#)]
46. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **2007**, *56*, 564–577. [[CrossRef](#)] [[PubMed](#)]
47. Zhang, D.; Gao, F.; Jakovlic, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)]
48. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **2017**, *34*, 772–773. [[CrossRef](#)]
49. Lanfear, R.; Calcott, B.; Kainer, D.; Mayer, C.; Stamatakis, A. Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* **2014**, *14*, 82. [[CrossRef](#)]
50. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
51. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574. [[CrossRef](#)]
52. Darling, A.E.; Mau, B.; Perna, N.T. Progressivemauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **2010**, *5*, e11147. [[CrossRef](#)]
53. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)] [[PubMed](#)]
54. Puttick, M.N. MCMCTreeR: Functions to prepare MCMCTree analyses and visualize posterior ages on trees. *Bioinformatics* **2019**, *35*, 5321–5322. [[CrossRef](#)] [[PubMed](#)]
55. Hao, S.; Xue, J.; Wang, Q.; Liu, Z. Yuguangia ordinata gen. et sp. nov., a new lycopsid from the Middle Devonian (Late Givetian) of Yunnan, China, and its phylogenetic implications. *Int. J. Plant Sci.* **2007**, *168*, 1161–1175. [[CrossRef](#)]
56. Banks, H.P.; Bonamo, P.M.; Grierson, J.D. *Leclercqia complexa* gen. et sp. nov., a new Lycopod from the late middle devonian of eastern New York. *Rev. Palaeobot. Palynol.* **1972**, *14*, 19–40. [[CrossRef](#)]
57. Bonamo, P.M. *Rellimia thomsonii* (*Progymnospermopsida*) from the middle devonian of New York State. *Am. J. Bot.* **1977**, *64*, 1272. [[CrossRef](#)]
58. Strother, P.K.; Traverse, A. Plant microfossils from llandoveryan and wenlockian rocks of Pennsylvania. *Palynology* **1979**, *3*, 1–21. [[CrossRef](#)]
59. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)]
60. Paradis, E.; Claude, J.; Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004**, *20*, 289–290. [[CrossRef](#)]
61. Yu, G.; Lam, T.T.Y.; Zhu, H.; Guan, Y. Two methods for mapping and visualizing associated data on phylogeny using GGTREE. *Mol. Biol. Evol.* **2018**, *35*, 3041–3043. [[CrossRef](#)]
62. Louca, S.; Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **2018**, *34*, 1053–1055. [[CrossRef](#)]
63. Robison, T.A.; Wolf, P.G. ReFernment: An R package for annotating RNA editing in plastid genomes. *Appl. Plant Sci.* **2019**, *7*, e01216. [[CrossRef](#)] [[PubMed](#)]
64. Gu, Y.; Yu, J.; Shang, H.; Liu, B.; Yan, Y. Complete chloroplast genome of *Isoetes hypsophila* (*Isoëtaceae*), the Endangered quillwort in China. *Mitochondrial DNA Part B Resour.* **2021**, *6*, 2908–2909. [[CrossRef](#)]
65. Palmer, J.D.; Thompson, W.F. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **1982**, *29*, 537–550. [[CrossRef](#)]
66. Hecht, J.; Grewe, F.; Knoop, V. Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of *Selaginella moellendorffii* mitochondria: The root of frequent plant mtDNA recombination in early tracheophytes. *Genome Biol. Evol.* **2011**, *3*, 344–358. [[CrossRef](#)] [[PubMed](#)]
67. Monroe, J.G.; Srikant, T.; Carbonell-Bejerano, P.; Becker, C.; Lensink, M.; Exposito-Alonso, M.; Klein, M.; Hildebrandt, J.; Neumann, M.; Kliebenstein, D.; et al. Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* **2022**, *602*, 101–105. [[CrossRef](#)] [[PubMed](#)]
68. Browne, P.D.; Nielsen, T.K.; Kot, W.; Aggerholm, A.; Gilbert, M.T.P.; Puetz, L.; Rasmussen, M.; Zervas, A.; Hansen, L.H. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *Gigascience* **2020**, *9*, g1aa008. [[CrossRef](#)]
69. Pace, N.R. Structure and synthesis of the ribosomal ribonucleic acid of prokaryotes. *Bacteriol. Rev.* **1973**, *37*, 562–603. [[CrossRef](#)]

70. Guo, W.; Grewe, F.; Mower, J.P. Variable frequency of plastid RNA editing among ferns and repeated loss of uridine-to-cytidine editing from vascular plants. *PLoS ONE* **2015**, *10*, e0117075. [[CrossRef](#)] [[PubMed](#)]
71. Oldenkott, B.; Yamaguchi, K.; Tsuji-Tsukinoki, S.; Knie, N.; Knoop, V. Chloroplast RNA editing going extreme: More than 3400 events of C-to-U editing in the chloroplast transcriptome of the lycophyte *Selaginella uncinata*. *Rna* **2014**, *20*, 1499–1506. [[CrossRef](#)]
72. Du, X.Y.; Lu, J.M.; Li, D.Z. Extreme plastid RNA editing may confound phylogenetic reconstruction: A case study of *Selaginella lycophytes*. *Plant Divers.* **2020**, *42*, 356–361. [[CrossRef](#)]
73. Tillich, M.; Lehwark, P.; Morton, B.R.; Maier, U.G. The evolution of chloroplast RNA editing. *Mol. Biol. Evol.* **2006**, *23*, 1912–1921. [[CrossRef](#)] [[PubMed](#)]
74. Smith, D.R. Unparalleled variation in RNA editing among selaginella plastomes. *Plant Physiol.* **2020**, *182*, 12–14. [[CrossRef](#)] [[PubMed](#)]
75. Rüdinger, M.; Volkmar, U.; Lenz, H.; Groth-Malonek, M.; Knoop, V. Nuclear DYW-type PPR gene families diversify with increasing RNA editing frequencies in liverwort and moss mitochondria. *J. Mol. Evol.* **2012**, *74*, 37–51. [[CrossRef](#)] [[PubMed](#)]
76. Kidston, R.; Lang, W.H. On Old Red Sandstone plants showing structure, from the Rhynie Chert Bed, Aberdeenshire. Part II. Additional notes on *Rhynia gwynne-vaughani*, Kidston and Lang: With descriptions of *Rhynia major*, N. Sp., and *Hornea lignieri*, N.G.N. sp. *Trans. R. Soc. Edinb.* **1920**, *52*, 603–627. [[CrossRef](#)]
77. Rickards, R.B. The age of the earliest club mosses: The Silurian Baragwanathia flora in Victoria, Australia. *Geol. Mag.* **2000**, *137*, 207–209. [[CrossRef](#)]
78. Evkaikina, A.I.; Berke, L.; Romanova, M.A.; Proux-Wéra, E.; Ivanova, A.N.; Rydin, C.; Pawlowski, K.; Voitsekhovskaja, O.V. The *Huperzia* selago shoot tip transcriptome sheds new light on the evolution of leaves. *Genome Biol. Evol.* **2017**, *9*, 2444–2460. [[CrossRef](#)] [[PubMed](#)]
79. Field, A.R.; Testo, W.; Bostock, P.D.; Holtum, J.A.M.; Waycott, M. Molecular phylogenetics and the morphology of the *Lycopodiaceae* subfamily *Huperzioidae* supports three genera: *Huperzia*, *Phlegmariurus* and *Phylloglossum*. *Mol. Phylogenet. Evol.* **2016**, *94*, 635–657. [[CrossRef](#)]