PNAS RESEARCH ARTICLE | PLANT BIOLOGY

Extraordinary preservation of gene collinearity over three hundred million years revealed in homosporous lycophytes

Cheng Li[a,1], David Wickell[b,c,1], Li-Yaung Kuo[d], Xueqing Chen[a], Bao Nie[a], Xuezhu Liao[a], Dan Peng[a], Jiaojiao Ji[a], Jerry Jenkins[e], Mellissa Williams[e], Shengqiang Shu[f], Christopher Plott[e], Kerrie Barry[f], Shanmugam Rajasekar[g], Jane Grimwood[e], Xiaoxu Han[a], Shichao Sun[a], Zhuangwei Hou[a], Weijun He[a], Guanhua Dai[h], Cheng Sun[i], Jeremy Schmutz[e,f], James H. Leebens-Mack[j], Fay-Wei Li[b,c,2], and Li Wang[a,k,2]

Edited by Loren Rieseberg, The University of British Columbia, Vancouver, BC, Canada; received July 31, 2023; accepted December 11, 2023

Homosporous lycophytes (Lycopodiaceae) are a deeply diverged lineage in the plant tree of life, having split from heterosporous lycophytes (*Selaginella* and *Isoetes*) ~400 Mya. Compared to the heterosporous lineage, Lycopodiaceae has markedly larger genome sizes and remains the last major plant clade for which no chromosome-level assembly has been available. Here, we present chromosomal genome assemblies for two homosporous lycophyte species, the allotetraploid *Huperzia asiatica* and the diploid *Diphasiastrum complanatum*. Remarkably, despite that the two species diverged ~350 Mya, around 30% of the genes are still in syntenic blocks. Furthermore, both genomes had undergone independent whole genome duplications, and the resulting intragenomic syntenies have likewise been preserved relatively well. Such slow genome evolution over deep time is in stark contrast to heterosporous lycophytes and is correlated with a decelerated rate of nucleotide substitution. Together, the genomes of *H. asiatica* and *D. complanatum* not only fill a crucial gap in the plant genomic landscape but also highlight a potentially meaningful genomic contrast between homosporous and heterosporous species.

homosporous lycophytes | genome evolution | whole genome duplication | gene collinearity | subgenome dominance

Lycophytes represent a vital resource for understanding the early evolution of vascular plants on land (1). They occupy an important phylogenetic position sister to all other vascular plants with a rich evolutionary history documented by fossil records dating back to the late Devonian (2). Lycophytes have independently evolved many complex traits alongside ferns and seed plants including photosynthetic leaves and heterospory.

Extant lycophytes comprise around 1,330 species in three deeply diverged families, Selaginellaceae, Isoetaceae, and Lycopodiaceae (3) (Fig. 1*A*). Like ferns, lycophytes also include both homosporous and heterosporous members, with Selaginellaceae and Isoetaceae being heterosporous and Lycopodiaceae being entirely homosporous (4). Homosporous plants produce only one type of spore, which develops into a bisexual gametophyte that produces both sperm and egg cells. On the other hand, heterosporous plants produce two types of spores (micro- and megaspore) that develop into separate male and female gametophytes. Homosporous ferns and lycophytes tend to have larger genome sizes than their heterosporous counterparts and are characterized by relatively high chromosome numbers (5).

Whole genome duplication (WGD) has long been recognized as a key driver of genome size evolution and species diversification and is relatively common among ferns and flowering plants (6–9). Interestingly in lycophytes, the history of WGDs varies substantially across the major lineages. No ancient WGDs have been detected in Selaginellaceae (10–12), and only one round of WGD has been conclusively demonstrated in Isoetaceae (13). Conversely, several independent ancient WGDs have been postulated in Lycopodiaceae based on transcriptomic data (9) but have yet to be verified with genome assemblies. In contrast to homosporous ferns and lycophytes, heterosporous plants tend to have relatively few chromosomes despite a similar history of WGD. In flowering plants, this is explained by the rapid loss and rearrangement of chromosomal material immediately following WGD during a process known as diploidization (14, 15). Thus, the large genomes and high chromosome counts in homosporous ferns and lycophytes might result from a distinct process of diploidization that proceeds via silencing and individual gene loss without the concomitant reduction in chromosome number seen in flowering plants (16, 17). A thorough characterization of WGDs in Lycopodiaceae, especially regarding genome evolution postduplication, is needed to clarify what underlies the genome size disparity between heterosporous and homosporous lycophytes.

In general, plant genomes represent a significant challenge to modern sequencing and assembly methods due to their large size, high repeat content, and pervasive history of

Significance

Lycophytes occupy a critical phylogenetic position sister to all other vascular plants. Unlike seed plants, they comprise heterosporous (Selaginellaceae and Isoetaceae) and homosporous (Lycopodiaceae) lineages. Homosporous plants have long been known to possess large genomes with considerably more chromosomes than heterosporous counterparts. However, limited genomic resources for homosporous lycophytes have hindered efforts to identify precise differences underlying this fundamental distinction. Here, we assembled chromosome-level genomes of homosporous lycophytes, *Huperzia asiatica* and *Diphasiastrum complanatum*. Despite 350 Mya of divergence and independent whole genome duplications, synteny is remarkably well preserved between these genomes. This, combined with significantly reduced nucleotide substitution rates, suggests a contrasting mode of genome evolution between heterosporous and homosporous lycophytes.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]C.L. and D.W. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: fl329@cornell.edu or wangli03@caas.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2312607121/-/DCSupplemental.

Published January 18, 2024.

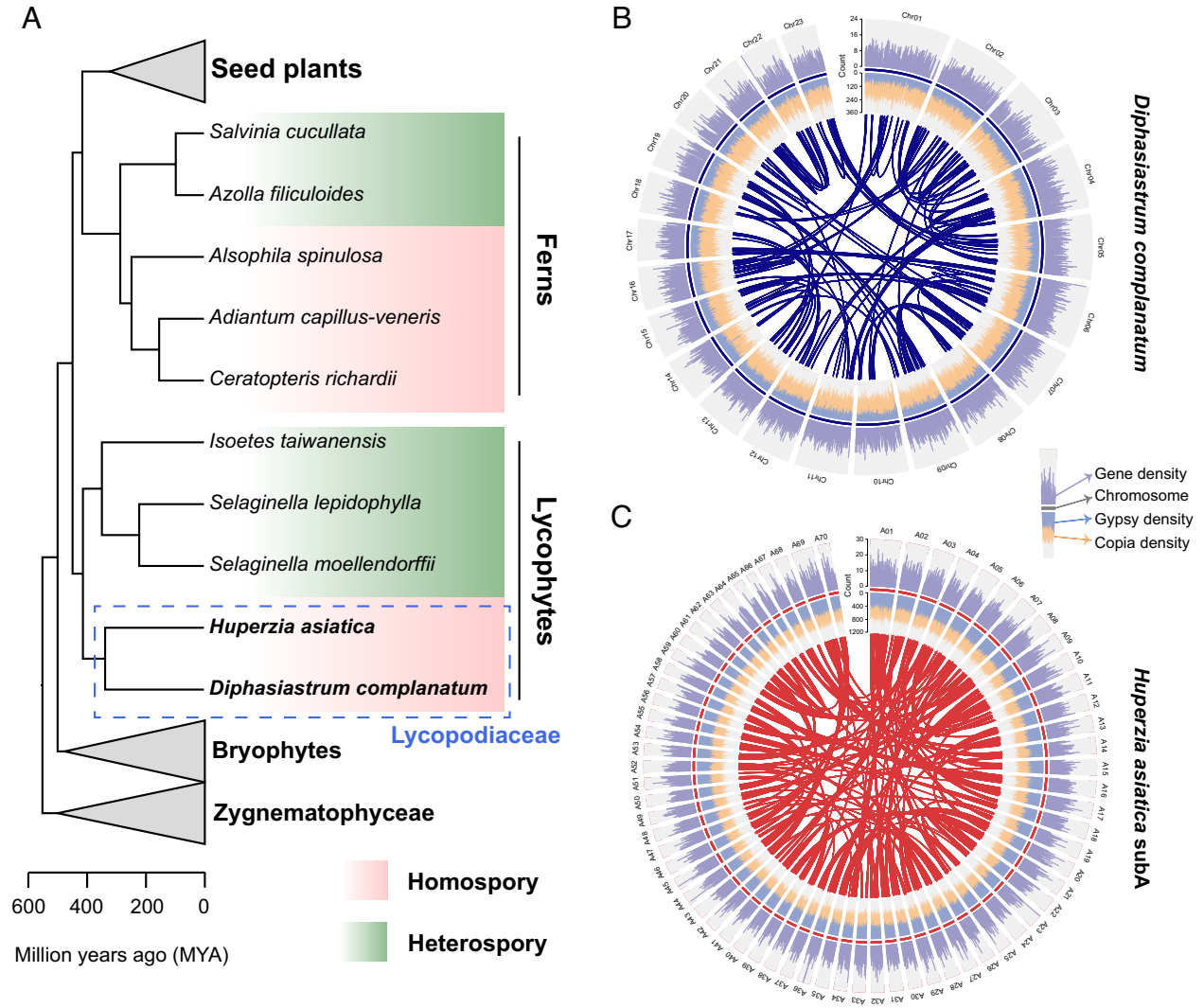

Fig. 1. The genomes of *H. asiatica* and *D. complanatum*. (*A*) Homosporous lycophytes (Lycopodiaceae), which diverged from the closest extant group over 400 Mya, are the last major lineage for which no genome had been available. (*B* and *C*) Genome features of *D. complanatum* and *H. asiatica* subA. The densities of genes and gypsy and copia LTRs for *D. complanatum* and *H. asiatica* subA were calculated with 500 Kb and 1 Mb nonoverlapped sliding windows, respectively. Intragenomic syntenies are displayed with connecting lines for *D. complanatum* and *H. asiatica* subA, respectively.

WGD. To date, most of the published lycophyte genomes came from heterosporous families Selaginellaceae and Isoetaceae which have relatively small genomes (10–13, 18–20). Given the deep divergences among these three lycophyte lineages (Fig. 1*A*), the lack of complete genome assemblies has left a gap in our knowledge of lycophyte genomics and hindered inferences of vascular plant evolution.

Here, we generated chromosome-level genome assemblies of the allotetraploid *Huperzia asiatica* and the diploid *Diphasiastrum complanatum*, belonging to Huperzioideae and Lycopodioideae subfamilies, respectively (*SI Appendix*, Fig. S1). Despite the divergence of the two subfamilies over 350 Mya (21), we observed a remarkable level of preserved synteny between these two genomes. While ancient intergenomic synteny has been previously described in other plants (22–24) and salmonid fish (25), the syntenic relationships recovered here are at least 100 My older than those reported in these other taxa. We also found a high degree of synteny within each genome following WGDs that occurred as much as 140 Mya. Further, we found little bias in genome fractionation and homoeologous gene expression in *H. asiatica* after the most recent allotetraploidization event. Such slow genome evolution appears to correlate with reduced rates of nucleotide substitutions compared to other lycophyte lineages. Our research fills a long-standing gap in land plant evolution and sheds light on the evolutionary history of early vascular plants tracing back hundreds of millions of years.

Results and Discussion

Genome Assembly and Annotation. Based on k-mer and flow cytometry analyses, the sizes of *H. asiatica* and *D. complanatum* genomes were estimated to be around 7.80 Gb and 1.60 Gb, respectively (*SI Appendix*, Figs. S2 and S3). Using a combination of PacBio CLR long reads, Illumina short reads, and Hi-C technology, we obtained chromosome-level assemblies for both genomes (*SI Appendix*, Fig. S4 and Dataset S1). For *H. asiatica*, the genome was assembled into 138 pseudochromosomes (n = 2X = 138) and 2,191 unanchored scaffolds, with a total length of 7.94 Gb and N50 of 57.02 Mb. For *D. complanatum*, the assembly consisted of 23 pseudochromosomes (n = X = 23) and 5,153 unanchored scaffolds, with a total length 1.74 Gb and N50 of 59.47 Mb. Both pseudochromosome counts align with literature reports (26–28).

Using subgenome-specific k-mers, we further partitioned the allotetraploid *Huperzia* genome into A and B subgenomes (hereafter referred to as "subA" and "subB") (*SI Appendix*, Figs. S5–S7). While we detected no evidence of recent large-scale homoeologous exchange between *H. asiatica* subgenomes based on k-mer sequences (*SI Appendix*, Fig. S6*B*), we could not exclude the possibility of ancient exchange shortly after allopolyploidization. The details on subgenome phasing are described in *SI Appendix*, *Supplementary Text*. The mapping rates of genomic Illumina reads were over 98.8% and 95.0% against the *H. asiatica* and *D. complanatum* assemblies (Dataset S2), respectively, indicating highly complete assemblies.

We found that 5.2% and 6.1% of the proteomes of *H. asiatica* and *D. complanatum* genomes were annotated as transcription-associated proteins (TAPs), which include transcription factors (TFs) and transcription regulators (TRs) (Dataset S7). *H. asiatica* and *D. complanatum* genomes contain a higher proportion of TAPs compared to bryophytes. The evolution of *YABBY*, a TF gene family that plays a key role in specifying leaf adaxial-abaxial polarity in seed plants (29–31), is particularly noteworthy. One *YABBY* ortholog was detected in *D. complanatum*, but none in *H. asiatica* despite the previous report of a single *YABBY* homolog in *H. selago* transcriptome data (32). We then searched transcriptomes of *H. miyoshiana*, *H. lucidula*, *H. javanica*, and *H. serrata*, and no *YABBY* homolog was identified. These results supported that *YABBY* may have been lost multiple times in *Huperzia* species, although the incompleteness in transcriptome data might also contribute to the spotty distribution. Moreover, taking account of previous reports that *YABBY* was absent in genomes of mosses (33), liverworts (34), *Isoetes* (13, 20), *Selaginella* (10–12), and ferns (35–38), but present in hornworts (39), the repeated loss of *YABBY* appears to suggest it is not necessary for abaxial/adaxial polarity in seed-free lineages. The dynamic evolution of *YABBY* is perplexing and warrants further research. Details on the evolution of gene families and transcription factors are described in *SI Appendix*, *Supplementary Text*.

Compared to the published heterosporous lycophytes [*Isoetes taiwanensis* (13), *Selaginella moellendorffii* (10), *S. lepidophylla* (12), and *S. tamariscina* (11)], our homosporous genomes comprised larger proportions of transposable elements (TEs) (81.86% for *H. asiatica* and 65.97% for *D. complanatum*), particularly long terminal repeats retrotransposons (LTR-RTs) (60.55% for *H. asiatica* and 53.68% for *D. complanatum*) (Dataset S3). A total of 82,725 and 31,430 high confidence protein-coding genes were annotated for *H. asiatica* and *D. complanatum*, respectively (Dataset S1). Both proteomes had high completeness scores from Benchmarking Universal Single-Copy Ortholog (BUSCO) with the "viridiplantae_odb10" database, indicating high annotation quality (90.4% for *H. asiatica*; 97.5% for *D. complanatum*) (Dataset S4). We found that *H. asiatica* and *D. complanatum* have 37.9-fold and 16.1-fold longer introns than the heterosporous lycophytes, respectively (Dataset S5), which is consistent with the positive correlation between intron length and genome size documented in other plant lineages (37, 38, 40). Compared to heterosporous lycophytes, the introns in homosporous lycophytes contain higher numbers of LTR (mean number 7.8 to 25.1 vs. 1.6 to 3.1) as well as larger total span (mean length 3.4 to 11.1 Kb vs. 0.5 to 1.9 Kb) (Dataset S6), suggesting that LTR insertion contributed to the intron length difference between homosporous and heterosporous genomes.

The chromosome-level assembly and annotation of lycophyte genomes allowed us to examine the distribution of repeats and genes in this lineage. In angiosperms, repeats and genes are generally unevenly distributed, with gene density increasing from centromere to telomere. However, we observe a different scenario in our genomes where the distribution of both appears to be homogeneous across the chromosomes (Fig. 1 *B* and *C* and *SI Appendix*, Fig. S8). A similar distribution was also reported in chromosome-level assemblies of bryophytes (33, 39) and ferns (35). Together, these results provide evidence that the genomic organization in seed-free plants might be quite distinct from seed plants.

Origin and Evolution of the Allotetraploid *H. asiatica*. To determine the parentage and timeline of subgenome divergence in *H. asiatica*, we reconstructed phylogenetic trees using genome and transcriptome data of *Huperzia* spp. and outgroup species (Fig. 2*A* and *SI Appendix*, Fig. S9). We found that *H. asiatica* subA and subB were clustered into two groups, which were named A-genome clade and B-genome clade, respectively. The divergence time between A- and B-genome clades was estimated to be ~35.3 Mya. *H. asiatica* subA and subB diverged from their closely related diploid species, *H. miyoshiana* and *H. lucidula*, ~17.9 Mya and ~19.5 Mya, respectively. In addition, the chloroplast phylogeny, which most likely tracks the maternal inheritance, further resolved that the A-genome clade is probably the maternal donor of *H. asiatica* (Fig. 2*A*).

Syntenic analysis of the two subgenomes of *H. asiatica* revealed 21,205 gene pairs, amounting to 55.87% of the total proteome in subA and 52.78% in subB, respectively, within 649 collinear blocks (*SI Appendix*, Fig. S10) suggesting limited chromosomal rearrangement since divergence of the progenitor lineage genomes ~35.3 Mya. Notably, only two large-scale chromosomal rearrangements (either chromosome fusion or fission events) were observed (Fig. 2 *B* and *C*), which was also confirmed by Hi-C (*SI Appendix*, Fig. S4*B*). One rearrangement involved B67 and its homoeologs, A67 and A68, and the other one involved B68 and its homoeologs, A69 and A70. Interestingly, we found that the average expression levels of genes within the rearranged chromosomes were significantly higher than the genomic background (*SI Appendix*, Fig. S11), which was not observed for the genes on the other chromosomes (*SI Appendix*, Fig. S12). Gene Ontology (GO) enrichment analysis uncovered that the genes residing in these rearranged chromosomes were significantly enriched in functions associated with diversification of specialized metabolism, such as terpenoid and tocopherol biosynthesis (*SI Appendix*, Fig. S13).

Although the *Huperzia* genus is widely distributed globally (https://www.gbif.org/species/2688450), *H. asiatica* is exclusively found in Changbai Mountain, located in the northeast of China (41). Consistent with its restricted distribution, we found that *H. asiatica* exhibited significantly lower genome-wide heterozygosity compared to closely related species (*SI Appendix*, Fig. S14), which resulted from its significantly smaller historical effective population size (Ne) (*SI Appendix*, Fig. S15).

Limited Subgenome Dominance in *H. asiatica*. Studies in flowering plants have shown that following allopolyploidization, one of the subgenomes often rose to dominance, which can be in the form of preferential retention of homoeologs and/or elevated gene expression levels (42–45). However, little is known about these processes outside of flowering plants. To compare the preferential gene retention after allopolyploidization, we characterized gene presence/absence variants (PAVs) between *H. asiatica* subgenomes. We found 3,742 (9.9% of the annotated genes in subA) and 4,570 (11.4% of the annotated genes in subB) orphan genes (i.e., lacking homoeologs) are present only in *H. asiatica* subA and subB, respectively. The average number of orphan genes per chromosome is 29.8 for subA and 31.0 for subB, and there is no significant difference between the subgenomes (*SI Appendix*, Fig. S16; Wilcoxon rank-sum test *P*-value = 0.52;

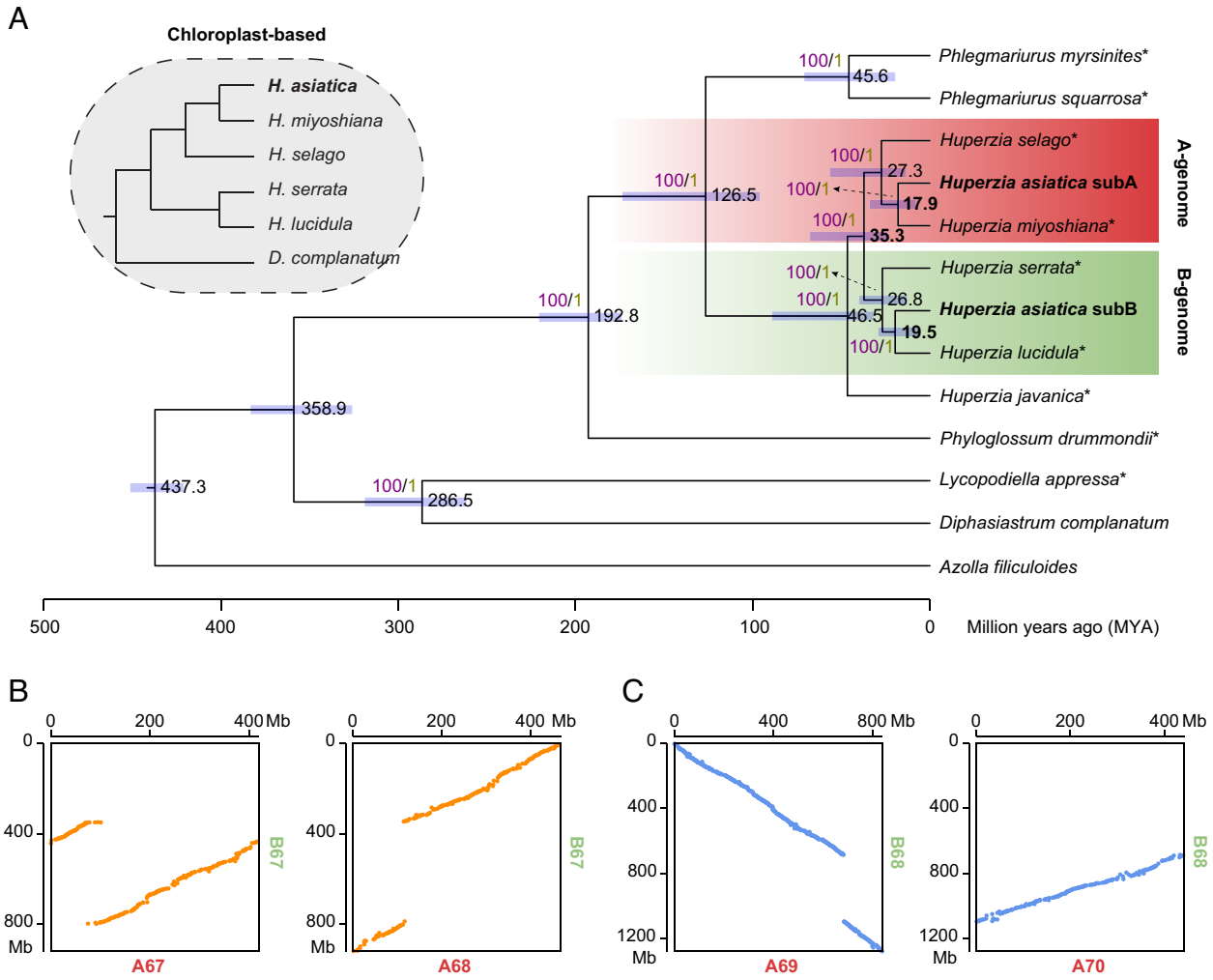

Fig. 2. Origin and evolution of the allotetraploid *H. asiatica*. (*A*) Species phylogeny of Lycopodiaceae inferred from nuclear 1,288 gene families. Numbers in purple and olive represent the bootstrap values of concatenation-based phylogeny and the ASTRAL support value of coalescence-based phylogeny, respectively. Numbers in black show the estimated divergence time. Blue bars at the nodes represent 95% CIs. Transcriptome-derived data are marked by an asterisk. The tree at *Upper Left* shows a plastome-based phylogeny from five *Huperzia* species and *D. complanatum*. (*B*) Synteny maps show that the chromosome fusion or fission event involved in B67 and its subA homoeologs, A67 and A68, in allotetraploid *Huperzia* genome. (*C*) Synteny maps show that the chromosome fusion or fission event involved in B68 and its subA homoeologs, A69 and A70, in allotetraploid *Huperzia* genome.

n = chromosome numbers), suggesting a pattern of nonpreferential retention of homoeologs. To quantify subgenome expression bias, we compared gene expression levels in 18,264 homoeologous gene pairs identified through synteny mapping (the light gray links in *SI Appendix*, Fig. S10 and Dataset S8). Generally, we found a weak expression divergence between homoeologs with average $\log_2$(subB/subA) expression ratios ranging from 0.02 to 0.07 in different tissue types (Fig. 3*A*). Gene pairs exhibiting homoeologous expression bias (HEB) were identified with cutoffs "*P*-value < 0.05 and | $\log_2$FoldChange | > 1." More than 70.0% of the 18,264 homoeolog pairs exhibited a nonbiased expression pattern (*SI Appendix*, Fig. S17). For pairs with significant HEB, the more highly expressed homoeolog was equally distributed between subA and subB (Fig. 3*B* and *SI Appendix*, Fig. S18). Overall, the two subgenomes exhibit a relatively balanced pattern of homoeolog gene expression. Homoeologous exchange has been hypothesized to mask the influence of genome dominance on gene expression (46), but here, we find little evidence of it (*SI Appendix*, Fig. S6*B*).

Next, we investigated possible factors that could give rise to the observed HEB. First, we explored the influence of LTR insertions on HEB gene pairs. We compared the LTR coverage in the flanking regions of the dominant homoeolog (with a higher expression) and the suppressed homoeolog (with a lower expression) and found that the dominant homoeologs had slightly but significantly lower LTR coverage than the suppressed counterparts (Fig. 3 *C* and *D*). This pattern was further corroborated in each of the sampled tissues (*SI Appendix*, Fig. S19). We then compared the ratio of nonsynonymous (Ka) to synonymous substitution rates (Ks) in homoeologous gene pairs to test whether pairs with HEB experienced different selective pressure. We found that gene pairs with HEB had a significantly higher Ka/Ks ratio than those without; however, when comparing subA- and subB-biased HEB gene pairs, the ratio does not significantly differ (Fig. 3*E*).

Taken together, our results suggest that neither subgenome has become dominant following allotetraploidization in *H. asiatica*. Relatively few homoeologous gene pairs show HEB and importantly, they do not consistently bias toward subA or subB. The age of the hybrid is obviously important in interpreting our findings. However, although we know that *H. asiatica* subgenomes diverged from the closest known diploid ~17.9 Mya, the hybridization and/or polyploidization event could take place anywhere between 17.9 Mya and the time we collected our sample. In other words, the lack of a clear subgenome dominance could be due to the recency

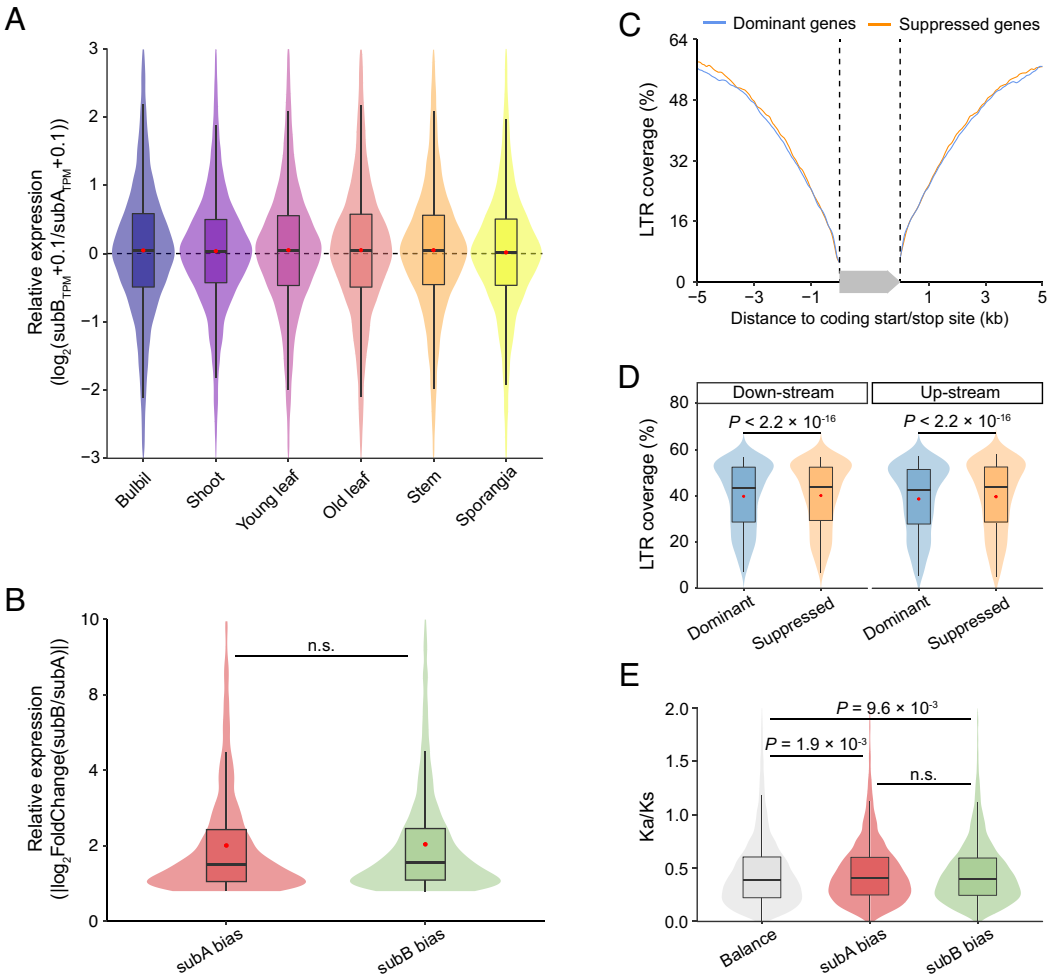

Fig. 3. Homoeolog expression bias between two subgenomes of *H. asiatica*. (*A*) Relative expression of homoeologous gene pairs in six different tissue types. (*B*) Comparisons for absolute value of $\log_2$FoldChange of homoeologous expression bias (HEB; cutoffs: *P*-value < 0.05 and | $\log_2$FoldChange | > 1) gene pairs between two subgenomes of allotetraploid *H. asiatica*. *P*-value was estimated by the Wilcoxon rank-sum test. "n.s" indicates no significant difference. (*C*) Different LTR coverage in the flanking regions between the dominant and the suppressed member for each homoeolog gene pair. The two dashed lines and a gray arrow indicate the gene region and orientation, respectively. (*D*) Comparisons for LTR coverage between the dominant and the suppressed member for each homoeologous gene pairs in up- and downstream of the coding region, respectively. *P*-value was estimated by the paired Wilcoxon test. (*E*) Comparisons for Ka/Ks values of homoeolog gene pairs with expression biased toward subA or subB or with expression balanced between the two subgenomes. *P*-value was estimated by the Wilcoxon rank-sum test.

of the hybrid or speaks to the slow evolutionary nature of *Huperzia* genomes (or the combination of both).

Highly Preserved Intergenomic Synteny Despite Deep Divergence. Comparing *H. asiatica* and *D. complanatum* genomes, we were surprised to find a large number of syntenic gene blocks. In total, 11,236 syntenic gene pairs in 1,515 blocks and 11,402 gene pairs in 1,566 blocks were identified between *D. complanatum* genome and *H. asiatica* subA and subB, respectively (Fig. 4*A* and *SI Appendix*, Fig. S20)—these amount to 26 to 36% of annotated genes in each genome. We can also find multiple *D. complanatum* chromosomes mapping strongly to eight *H. asiatica* chromosomes (four from each subgenome; Fig. 4*B*).

To compare our results with other vascular plants, we assessed synteny between flowering plants *Amborella trichopoda* (22), *Prunus persica* (47), *Vitis vinifera* (48), and *Theobroma cacao* (49), between homosporous ferns, *Alsophila spinulosa* (35), *Adiantum capillus-veneris* (37), and *Ceratopteris richardii* (38), and between heterosporous pteridophytes, *Azolla filiculoides* (36), *Isoetes taiwanensis* (13), *Isoetes sinensis* (20), and *Selaginella kraussiana* (19) (*SI Appendix*, Figs. S21–S23). Among flowering plants, we found that 27 to 32% of genes were retained in synteny between *Amborella trichopoda* and the other three angiosperms (*SI Appendix*, Fig. S21). The range for homosporous ferns is much greater with 5 to 32% of genes exhibiting collinearity (*SI Appendix*, Fig. S22). While the proportions of collinear genes broadly overlap with those between *D. complanatum* and *H. asiatica*, the age of the collinear relationships preserved in Lycopodiaceae is truly unprecedented given that *H. asiatica* and *D. complanatum* diverged roughly 368 Mya (21). As a comparison, *Amborella trichopoda* diverged from other flowering plants between 197.5 to 246.5 Mya (50), and the homosporous ferns used in our analysis diverged around 246 Mya (*SI Appendix*, Fig. S36). Finally, we found very little synteny between recently published chromosome-level assemblies for heterosporous lycophytes *I. sinensis* and *S. kraussiana* (*SI Appendix*, Fig. S23). Though they are thought to have diverged around the same time as *Huperzia* and *Diphasiastrum*, they retained only 4.4 to 7.2% of genes in synteny.

Highly Preserved Intragenomic Synteny Following Ancient WGDs. The preservation of synteny can also be observed within each *D. complanatum* and *H. asiatica* genome. In *D. complanatum*, we found 1,422 gene pairs contained in 216 collinear blocks accounting for 10.8% of annotated genes (Fig. 1*B* and *SI Appendix*,

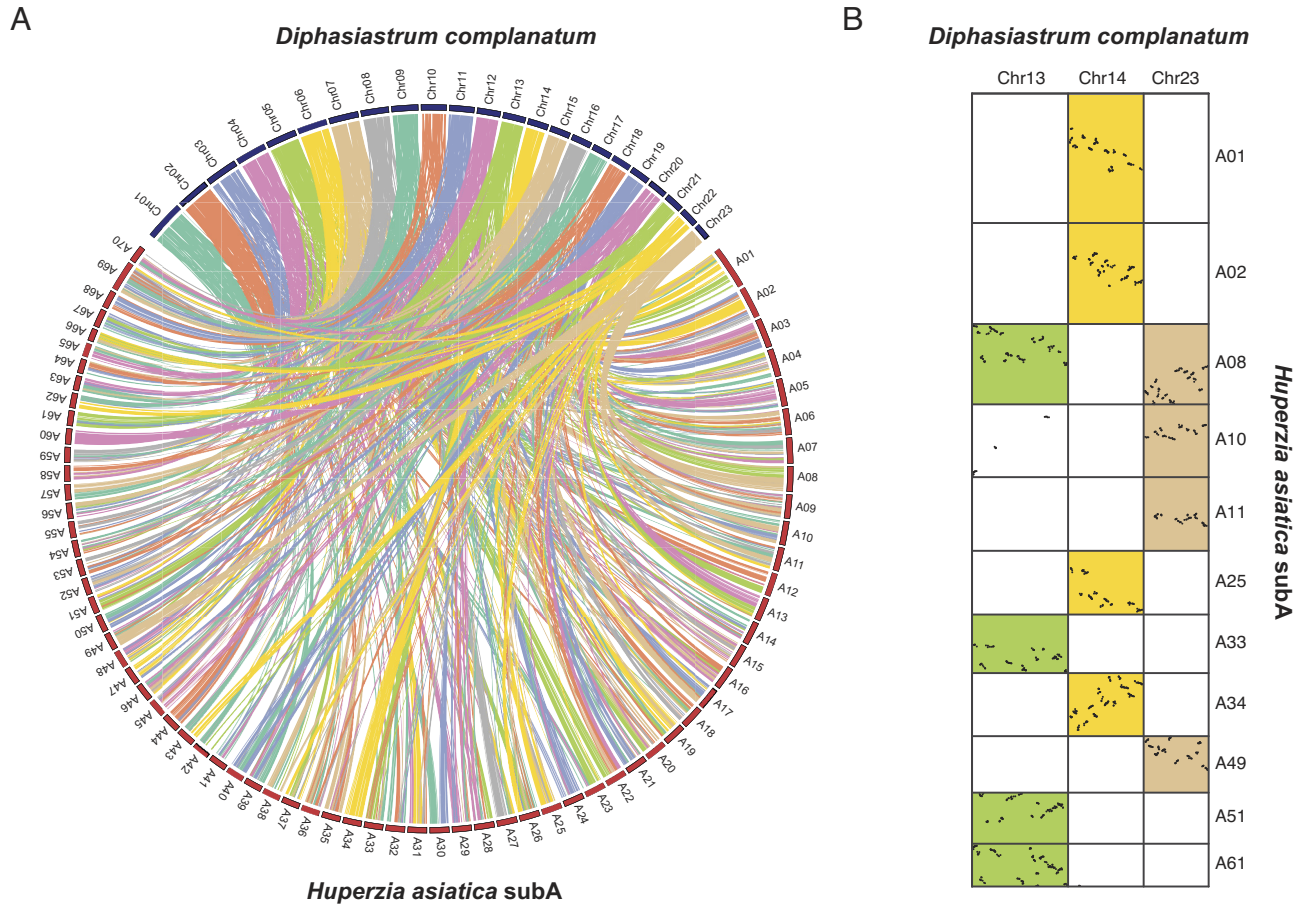

Fig. 4. Preserved synteny between distantly related *H. asiatica* and *D. complanatum*. (*A*) Syntenic relationship between *H. asiatica* subA and *D. complanatum* genomes. Collinear gene blocks are connected by ribbons. (*B*) Syntenic dotplot of representative chromosomes that show a clear "1:4" relationship comparing *D. complanatum* to *H. asiatica* subA.

Fig. S24). In *H. asiatica* subA and subB, we identified 3,284 gene pairs in 338 blocks (17.3% of annotated genes) and 3,778 gene pairs in 429 blocks (18.8% of annotated genes), respectively (Fig.1*C* and *SI Appendix*, Figs. S8 and S24). The intragenomic synteny in *D. complanatum* and *H. asiatica* genomes is most likely the result of ancient WGDs. Syntenic depth analysis found a 2:4 relationship between collinear blocks of genes in *D. complanatum* and either of the *H. asiatica* subgenomes (*SI Appendix*, Fig. S20), suggesting two WGDs (termed "Huper-α" and Huper-β") predating the divergence of *H. asiatica* subgenomes and a single WGD (termed "Lyco-α") in the ancestry of *D. complanatum*.

To further place these WGD events onto the phylogeny, we used both synonymous substitutions per site (Ks) and gene tree-species tree reconciliation approaches. Ks analysis of paralogs in *D. complanatum* produced a plot with a single peak centered around Ks = 0.45, suggesting that the Lyco-α event may have occurred following its divergence from *Huperzia* (Fig. 5*A* and *SI Appendix*, Fig. S25). For *H. asiatica*, the Ks plot for each subgenome has a single prominent peak near Ks = 0.2 (Fig. 5*B* and *SI Appendix*, Fig. S26). Although we might expect to see two peaks corresponding to the Huper-α and Huper-β events, given the extraordinarily low orthologous divergence observed between *Huperzia* and *Diphasiastrum* (Fig. 5 *A* and *B*), it is possible that two WGDs occurring close together would produce overlapping Ks distributions. This hypothesis seems to be supported by a Ks plot restricted to syntenic gene pairs that yields a single peak in each subgenome distributed around Ks = 0.26 (Fig. 5*B*). Our Ks analyses on related *Huperzia* species revealed an even more complex history of recurrent and independent WGDs throughout the family, a pattern consistent with the high and variable chromosome counts reported. Because of the frequent duplications in this lineage, it is impossible to place WGD events based solely on Ks.

Using the Multi-tAxon Paleopolyploid Search (MAPS) algorithm, our gene tree-species tree reconciliation approach corroborated that Lyco-α, Huper-α, and Huper-β are all independent events. We found no support for an earlier, shared duplication in the common ancestor of Lycopodiaceae. A MAPS analysis focused on Lycopodioideae (the subfamily including *Diphasiastrum*) placed Lyco-α prior to the divergence of *Dendrolycopodium* and *Diphasiastrum* (Fig. 5*C*). A separate analysis focused on the Huperzioideae implicated Huper-α and Huper-β occurred following the divergence of *Huperzia* from its sister genus *Phlegmariurus* (Fig. 5*D* and *SI Appendix*, Fig. S27).

Based on their phylogenetic placements, the three WGD events we uncovered here are all ancient (*SI Appendix*, Fig. S28). Lyco-α in particular is likely at least 139 Mya, given it predates the divergence between *Diphasiastrum* and *Dendrolycopodium* (21). Despite the antiquity of this duplication, 10.8% of the annotated genes are retained in synteny. Though this is roughly half the proportion of genes that exhibit syntenic relationships in the homosporous fern *Alsophila spinulosa* (*SI Appendix*, Fig. S29), it is still far greater than what has been reported for heterosporous lineages following ancient WGD events. For instance, we recovered just 605 collinear genes representing 3.2% of the genome in the heterosporous fern *Azolla filiculoides* from a WGD that occurred roughly 100 Mya

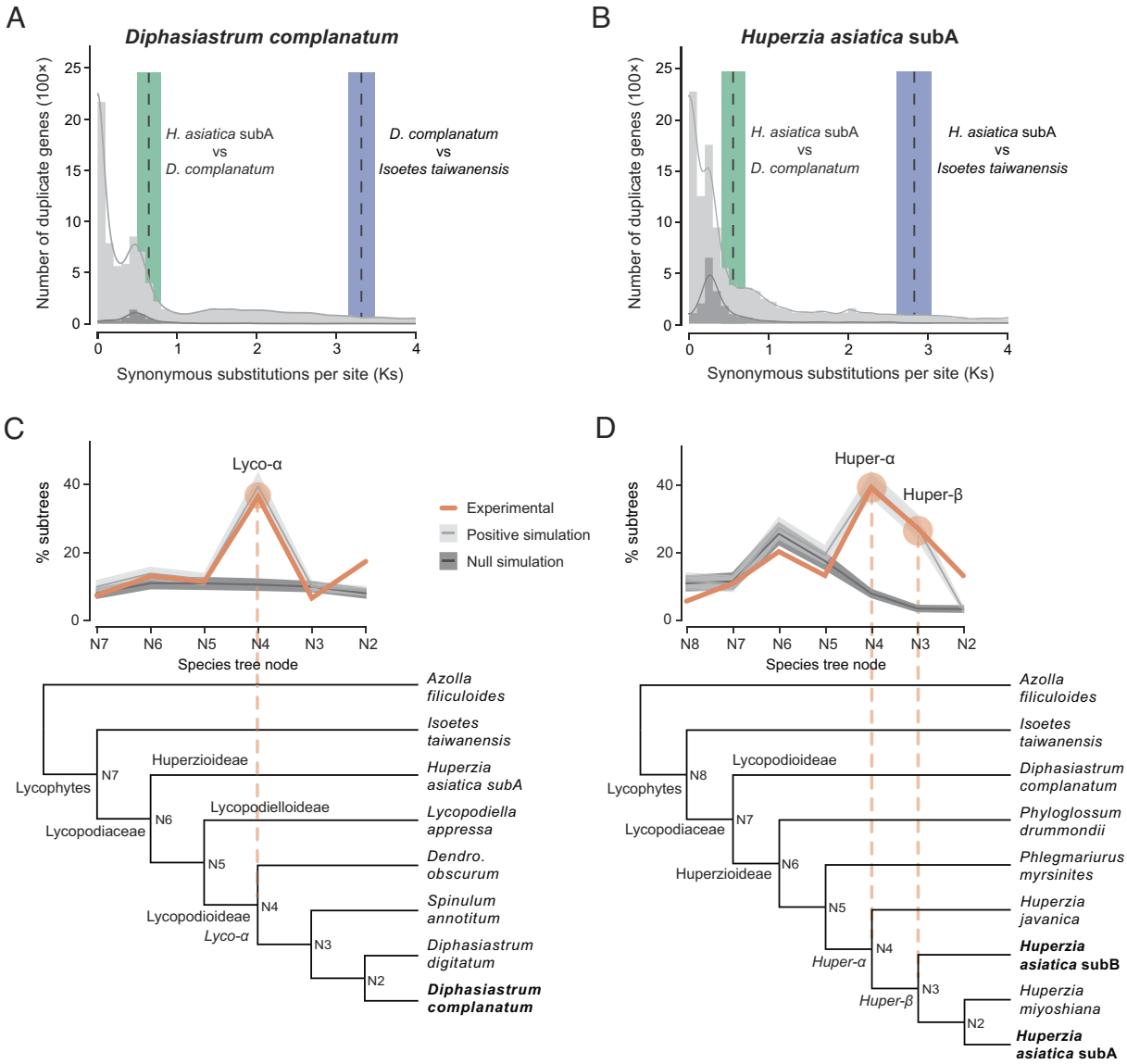

Fig. 5. Evidence for ancient WGD in *H. asiatica* and *D. complanatum*. (*A* and *B*) Ks plots of paralogous genes in *D. complanatum* (*A*) and *H. asiatica* subA (*B*). Analysis including all paralogs was shown in light gray, and analysis restricted to collinear gene pairs was shown in dark gray. Dashed vertical lines indicate mean adjusted orthologous divergence for species indicated; shaded areas represent one SD. (*C* and *D*) Summary of MAPS analyses of *D. complanatum* (*C*) and *H. asiatica* subA (*D*). Dark lines in the center of the shaded regions represent average values for null and positive gene tree simulations; shaded areas around each line show the SD for gene tree simulations. Dashed orange lines mark the inferred WGDs.

(*SI Appendix*, Fig. S30) (36, 51). Among flowering plants, the ancestral angiosperm WGD detected in the *Amborella trichopoda* genome is represented by only 1.8% of annotated genes (*SI Appendix*, Fig. S31). That being said, a more appropriate comparison exists between *D. complanatum* and the heterosporous lycophyte *I. taiwanensis* (13). Though the WGD in *Isoetes* has not been dated with certainty, a mere 3.8% of genes were found to possess syntenic relationships in *I. taiwanensis* (*SI Appendix*, Fig. S32). This relatively slow rate of genome rearrangement and fractionation within Lycopodiaceae echoes our finding on intergenomic synteny between *D. complanatum* and *H. asiatica* reported above.

Reciprocal Fractionation May Obscure Intragenomic Synteny in Homosporous Lineages. Despite a predominant 2:4 relationship between syntenic blocks of genes in *D. complanatum* and *H. asiatica* subgenomes (*SI Appendix*, Fig. S20), we found relatively few genes showing the expected 2:2 relationship within the *H. asiatica* subgenome (*SI Appendix*, Fig. S24). While this may at first seem counterintuitive, it can be explained by the process of differential fractionation of genes in each species following WGD. Immediately following duplication, species begin to undergo the process of diploidization, by which diploid patterns of gene expression and inheritance are restored through gene fractionation and chromosomal rearrangement. In heterosporous species, these processes are tightly coupled with many genes being lost rapidly via illegitimate recombination (15). However, high numbers of chromosomes in homosporous lineages of ferns and lycophytes suggest that such rapid, large-scale losses and rearrangements occur less frequently (52, 53). Instead, the primary processes driving diploidization in homosporous plants likely involve silencing and eventual loss of individual genes (52, 54). Thus, through the reciprocal loss of paralogs along homoeologous chromosomes, it is possible to produce two sets of homoeologous chromosomes with a relatively high degree of intergenomic synteny despite exhibiting little to no intragenomic synteny (Fig. 6). If this

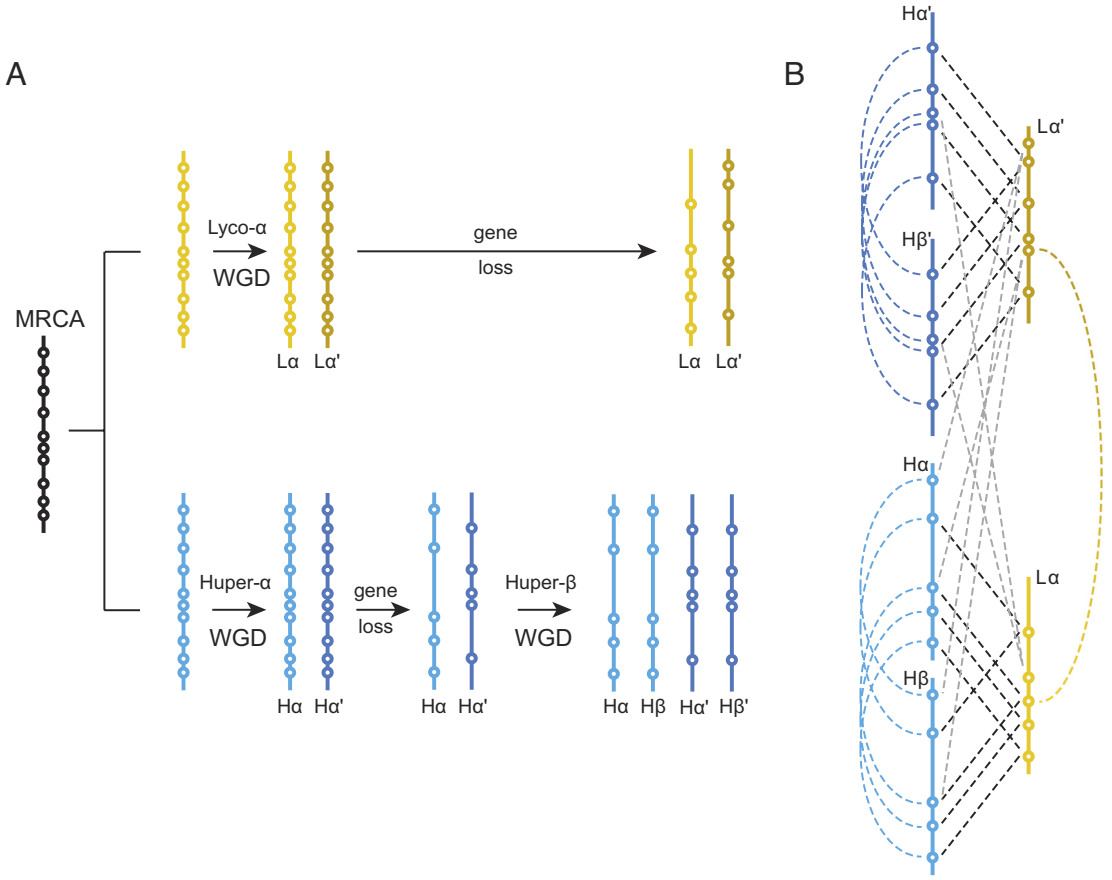

Fig. 6. Diagram of hypothetical syntenic relationships following multiple rounds of WGD and unequal gene fractionation. An illustration showing (*A*) a hypothetical scenario that could potentially preserve a "2:4" pattern of interspecific synteny between two species while obscuring (*B*), intraspecific syntenic relationships within them, similar to what is observed between and within the genomes of *H. asiatica* and *D. complanatum*.

initial gene loss was followed by a second WGD in *H. asiatica*, we would then expect a 1:1 relationship to be restored between newly duplicated chromosomes assuming that there has been less fractionation following this more recent WGD. In fact, when we zoom in on scaffolds that exhibit a 2:4 relationship between *Huperzia* and *Diphasiastrum*, this is precisely what we see (*SI Appendix*, Figs. S33–S35). Large overlapping blocks are anchored by distinct genes scattered across the same region of the scaffold with relatively low levels of intragenomic synteny relative to intergenomic synteny. Despite this process of diploidization by fractionation, the retention of long-range synteny both within and between these species is remarkable given the timescales involved and unlike anything previously described in heterosporous plants. Furthermore, this pattern seems to corroborate the hypothesis that diploidization in homosporous lineages is largely driven by silencing and fractionation of individual genes rather than large-scale structural changes (52, 53).

Highly Preserved Synteny Is Linked to Decelerated Rates of Nucleotide Substitution. The collinearity found in *H. asiatica* and *D. complanatum*, both within and between genomes, are in stark contrast to sister, heterosporous lineages [*Isoetes* (13, 20) and *Selaginella* (19)]. Our placement of the Lyco-α WGD event prior to the divergence of *Diphasiastrum* and *Dendrolycopodium* renders the age of intraspecific synteny within *D. complanatum* to be on par with that recently found in tree fern *Alsophila spinulosa* (35). Similarly, while limited examples of ancient inter- and intraspecific collinearity have been reported in species of fish (25, 55) and moss (23, 24), divergence time estimates for *Diphasiastrum* and *Huperzia* are considerably older (21). Surprisingly, given their great evolutionary distance, orthologous divergence between *H. asiatica* and *D. complanatum* is relatively low (Ks ~ 0.5 to 0.7; Fig. 5 *A* and *B*) implying a slow rate of substitution. As previously described for tree ferns (35), we found decelerated rates of substitution in Lycopodiaceae compared to heterosporous lycophytes *Selaginella* and *Isoetes*. Most of the gene families examined show signs of deceleration and 22 to 31% of the comparisons are significant (*SI Appendix*, Fig. S36). These results, combined with similar findings in *Alsophila spinulosa* (35), indicate that the gradual processes of chromosomal rearrangement and fractionation may be correlated with relatively slow rates of substitution in homosporous seed-free lineages.

Putative Genomic Difference between Homosporous and Heterosporous Lineages. Our study suggests that homosporous lycophytes likely have a contrasting mode of genome evolution compared to heterosporous lycophytes and sheds light on the process of diploidization following WGD. Homosporous lineages are known to harbor exceptionally large genomes with high chromosome counts relative to other vascular plants. This contrast has puzzled botanists for over half the century (4, 5, 53, 56, 57), and no mechanistic explanation has been widely accepted so far. Our research reveals a number of factors contributing to the markedly larger genome sizes of homosporous lycophytes, including an abundance of TE insertions, extreme intron length, and multiple rounds of WGD. Importantly, we found that the rates of molecular evolution decelerated, specifically those of substitution, fractionation, and chromosomal rearrangement.

It is tempting to generalize our findings to all homosporous vascular plants, especially considering that similar patterns of retained synteny and reduced substitution rates have been reported in the homosporous tree fern *Alsophila spinulosa* (35). However, the more recent publications of *Ceratopteris richardii* (38) and *Adiantum capillus-veneris* (37) genomes, both from Pteridaceae, suggest the same might not be true in other homosporous ferns. Both species exhibited relatively little intragenomic synteny despite experiencing a shared WGD at roughly the same time as *Alsophila spinulosa*. In addition, both were previously found to have significantly higher substitution rates than *Alsophila spinulosa* (35). As such, the effects of WGD, subsequent diploidization, and evolutionary rates on the size of homosporous fern genomes may be more nuanced and await further study. While the current sampling of homosporous genomes is too narrow to give us a complete picture, our study has nevertheless identified a possible genomic contrast between homosporous and heterosporous species. This testable hypothesis will set the stage for future comparative studies on seed-free genomes.

Materials and Methods

The *H. asiatica* and *Diphasiastrum complanatum* plants were collected from China and Taiwan. We sequenced *H. asiatica* and *D. complanatum* using a combination of sequencing technologies, including long-read sequencing from Pacific Biosciences (PacBio), Illumina short-read sequencing, and chromosome conformation capture using Hi-C sequencing. De novo assemblies for *H. asiatica* and *D. complanatum* were generated using CANU v2.0 (58) and MECAT2 (59), respectively. HiC scaffolding was performed by 3D-DNA pipeline v180922 (60). Phylogenetic trees were constructed using low-copy orthologous genes with RAxML v8.2.12 (61) and ASTRAL v5.7.1 (62), respectively. Synteny blocks were identified using the python implementation of MCSCAN v0.84 (63). Genome-wide heterozygosity was estimated using ANGSD v0.935 (64) based on a Site Frequency Spectrum (SFS). The Pairwise Sequentially Markovian Coalescent (PSMC) method (65) was used to infer the population size history of *Huperzia* species and *D. complanatum*. Homoeologous expression bias (HEB) gene sets in *H. asiatica* were identified between all the homoeologous gene pairs of two subgenomes using the DESeq2 v3.17 package (66). Paralogous Ks estimation and fitting of generalized mixture models were conducted with the WGD package v1.1.2 (67). We used MAPS (68) to place WGD events onto the phylogeny. Substitution rates across lycophytes were estimated with protein-coding genes by using PAML v4.4 (69). Details on the materials and methods used in this study can be found in *SI Appendix*.

Data, Materials, and Software Availability. The raw data of genome and transcriptome sequencing of *H. asiatica* have been deposited to the Genome Sequence Archive at the National Genomics Data Center (NGDC) under BioProject No. PRJCA013778 (70). The genome assemblies and annotations of *H. asiatica* and all the chloroplast genomes assembled are available at figshare platform (https://figshare.com/projects/Huperzia_asiatica_genome/169145) (71). The raw data of genome and transcriptome sequencing of *D. complanatum* have been deposited in the NCBI SRA under BioProject No. PRJNA914350 (72). The genome assemblies and annotations of *D. complanatum* can be found in Phytozome (https://phytozome-next.jgi.doe.gov/info/Dcomplanatum_v3_1) (73).

ACKNOWLEDGMENTS. This study was supported by the National Key Research and Development Program of China (Grant No. 2023YFA0915802), National Natural Science Foundation of China (Grant Nos. 32070242 and 32300223), Innovation Program of Chinese Academy of Agricultural Sciences, Key project at central government level: the ability establishment of sustainable use for valuable Chinese medicine resources (Grant No. 2060302), Shenzhen Science and Technology Program (Grant No. KQTD20161130104826 51), Special funds for Science Technology Innovation and Industrial Development of Shenzhen Dapeng New District (Grant Nos. RC201901-05 and PT201901-19), and Shenzhen Fundamental Research Program grant (Grant No. 20220817165436004). The work (proposal: 10.46936/10.25585/60001405) also conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.

Author affiliations: [a]Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Key Laboratory of Synthetic Biology, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518000, China; [b]Boyce Thompson Institute, Ithaca, NY 14853; [c]Plant Biology Section, Cornell University, Ithaca, NY 14853; [d]Institute of Molecular and Cellular Biology, National Tsing Hua University, Hsinchu 300044, Taiwan; [e]Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806; [f]United States Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [g]Arizona Genomics Institute, School of Plant Sciences, University of Arizona, Tucson, AZ 85721; [h]Research Station of Changbai Mountain Forest Ecosystems, Chinese Academy of Sciences, Yanji 133000, China; [i]College of Life Sciences, Capital Normal University, Beijing 100048, China; [j]Department of Plant Biology, University of Georgia, Athens, GA 30602; and [k]State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-di Herbs, Beijing 100700, China

Author contributions: F.-W.L. and L.W. designed research; C.L., L.-Y.K., X.C., B.N., J. Jiao, X.H., S. Sun, G.D., and C.S. collected plant samples and performed experiments; C.L., D.W., X.L., D.P., J. Jenkins, M.W., S. Shu, C.P., K.B., S.R., J.G., Z.H., W.H., J.S., and J.H.L.-M. analyzed data; and C.L., D.W., F.-W.L., and L.W. wrote the paper.

1. V. Spencer, Z. Nemec Venza, C. J. Harrison, What can lycophytes teach us about plant evolution and development? Modern perspectives on an ancient lineage. *Evol. Dev.* **23**, 174–196 (2020).
2. P. Kenrick, P. R. Crane, The origin and early evolution of plants on land. *Nature* **389**, 33–39 (1997).
3. E. Schuettpelz *et al.*, A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* **54**, 563–603 (2016).
4. S. P. Kinosian, C. A. Rowe, P. G. Wolf, Why do heterosporous plants have so few chromosomes? *Front. Plant. Sci.* **13**, 807302 (2022).
5. F.-W. Li, The chromosome hoarding syndrome of (some) ferns and lycophytes. *Nat. Rev. Genet.* **24**, 737–737 (2023).
6. T. E. Wood *et al.*, The frequency of polyploid speciation in vascular plants. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13875–13879 (2009).
7. P. S. Soltis, D. E. Soltis, Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Plant Biol.* **30**, 159–165 (2016).
8. S. Wu, B. Han, Y. Jiao, Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. *Mol. Plant* **13**, 59–71 (2020).
9. J. H. Leebens-Mack *et al.*, One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679–685 (2019).
10. J. A. Banks *et al.*, The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**, 960–963 (2011).
11. Z. Xu *et al.*, Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. *Mol. Plant* **11**, 983–994 (2018).
12. R. VanBuren *et al.*, Extreme haplotype variation in the desiccation-tolerant clubmoss *Selaginella lepidophylla*. *Nat. Commun.* **9**, 13 (2018).
13. D. Wickell *et al.*, Underwater CAM photosynthesis elucidated by *Isoetes* genome. *Nat. Commun.* **12**, 6348 (2021).
14. M. Chester *et al.*, Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae). *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1176–1181 (2012).
15. Z. Li *et al.*, Patterns and processes of diploidization in land plants. *Annu. Rev. Plant Biol.* **72**, 387–410 (2021).
16. J. Clark *et al.*, Genome evolution of ferns: Evidence for relative stasis of genome size across the fern phylogeny. *New Phytol.* **210**, 1072–1082 (2016).
17. C. H. Haufler, Ever since Klekowski: Testing a set of radical hypotheses revives the genetics of ferns and lycophytes. *Am. J. Bot.* **101**, 2036–2042 (2014).
18. J.-G. Yu *et al.*, The first homosporous lycophyte genome revealed the association between the recent dynamic accumulation of LTR-RTs and genome size variation. *Plant Mol. Biol.* **112**, 325–340 (2023).
19. W. Liu *et al.*, Genome and transcriptome of *Selaginella kraussiana* reveal evolution of root apical meristems in vascular plants. *Curr. Biol.* **33**, 4085–4097 (2023).
20. J. Cui *et al.*, Chromosome-level reference genome of tetraploid *Isoetes sinensis* provides insights into evolution and adaption of lycophytes. *GigaScience* **12**, 1–15 (2022).
21. W. Testo, A. Field, D. Barrington, Overcoming among-lineage rate heterogeneity to infer the divergence times and biogeography of the clubmoss family Lycopodiaceae. *J. Biogeogr.* **45**, 1929–1941 (2018).
22. V. A. Albert *et al.*, The *Amborella* genome and the evolution of flowering plants. *Science* **342**, 1241089 (2013).
23. J. Yu *et al.*, Chromosome-level genome assemblies of two Hypnales (mosses) reveal high intergeneric synteny. *Genome Biol. Evol.* **14**, evac020 (2022).
24. S. B. Carey *et al.*, Gene-rich UV sex chromosomes harbor conserved regulators of sexual development. *Sci. Adv.* **7**, eabh2488 (2021).
25. T. Sävilammi *et al.*, The chromosome-level genome assembly of European grayling reveals aspects of a unique genome evolution process within salmonids. *G3-Genes Genom. Genet.* **9**, 1283–1294 (2019).
26. J. Beitel, F. Wagner, The chromosomes of *Lycopodium lucidulum*. *Am. Fern J.* **72**, 33–35 (1982).
27. J. Holub, *Diphasiastrum*, a new genus in Lycopodiaceae. *Proslia Praha* **47**, 97–110 (1975).
28. A. Rice *et al.*, The Chromosome Counts Database (CCDB)–a community resource of plant chromosome numbers. *New Phytol.* **206**, 19–26 (2015).

29. F. Du, C. Guan, Y. Jiao, Molecular mechanisms of leaf morphogenesis. *Mol. Plant* **11**, 1117–1134 (2018).
30. R. Sarojam *et al.*, Differentiating arabidopsis shoots from leaves by combined YABBY activities. *Plant Cell* **22**, 2113–2130 (2010).
31. T. Yamaguchi *et al.*, The *YABBY* gene *DROOPING LEAF* regulates carpel specification and midrib development in Oryza sativa. *Plant Cell* **16**, 500–509 (2004).
32. A. I. Evkaikina *et al.*, The *Huperzia selago* shoot tip transcriptome sheds new light on the evolution of leaves. *Genome Biol. Evol.* **9**, 2444–2460 (2017).
33. D. Lang *et al.*, The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* **93**, 515–533 (2018).
34. J. L. Bowman *et al.*, Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* **171**, 287–304 (2017).
35. X. Huang *et al.*, The flying spider-monkey tree fern genome provides insights into fern evolution and arborescence. *Nat. Plants* **8**, 500–512 (2022).
36. F. W. Li *et al.*, Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nat. Plants* **4**, 460–472 (2018).
37. Y. Fang *et al.*, The genome of homosporous maidenhair fern sheds light on the euphyllophyte evolution and defences. *Nat. Plants* **8**, 1024–1037 (2022).
38. D. B. Marchant *et al.*, Dynamic genome evolution in a model fern. *Nat. Plants* **8**, 1038–1051 (2022).
39. F. W. Li *et al.*, *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat. Plants* **6**, 259–272 (2020).
40. S. Niu *et al.*, The Chinese pine genome and methylome unveil key features of conifer evolution. *Cell* **185**, 204–217 (2022).
41. N. Shrestha, X. C. Zhang, On the presence of North American clubmoss *Huperzia lucidula* (Lycopodiaceae) in China: An intercontinental disjunction or misidentification. *Phytotaxa* **219**, 243–252 (2015).
42. S. P. Otto, The evolutionary consequences of polyploidy. *Cell* **131**, 452–462 (2007).
43. B. McClintock, The significance of responses of the genome to challenge. *Science* **226**, 792–801 (1984).
44. P. P. Edger *et al.*, Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower. *Plant Cell* **29**, 2150–2167 (2017).
45. K. A. Bird *et al.*, Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. *New Phytol.* **230**, 354–371 (2021).
46. K. A. Bird, R. VanBuren, J. R. Puzey, P. P. Edger, The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytol.* **220**, 87–93 (2018).
47. I. Verde *et al.*, The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet.* **45**, 487–494 (2013).
48. O. Jaillon *et al.*, The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–467 (2007).
49. J. C. Motamayor *et al.*, The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biol.* **14**, r53 (2013).
50. J. L. Morris *et al.*, The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2274–E2283 (2018).
51. J. A. Pelosi, E. H. Kim, W. B. Barbazuk, E. B. Sessa, Phylotranscriptomics illuminates the placement of whole genome duplications and gene retention in ferns. *Front. Plant Sci.* **13**, 882441 (2022).
52. C. H. Haufler, Electrophoresis is modifying our concepts of evolution in homosporous pteridophytes. *Am. J. Bot.* **74**, 953–966 (1987).
53. M. S. Barker, P. G. Wolf, Unfurling fern biology in the genomics age. *BioScience* **60**, 177–185 (2010).
54. E. Pichersky, D. Soltis, P. Soltis, Defective chlorophyll a/b-binding protein genes in the genome of a homosporous fern. *Proc. Natl. Acad. Sci. USA* **87**, 195–199 (1990).
55. T. J. Krabbenhoft *et al.*, Chromosome-level genome assembly of Chinese sucker (*Myxocyprinus asiaticus*) reveals strongly conserved synteny following a catostomid-specific whole-genome duplication. *Genome Biol. Evol.* **13**, evab190 (2021).
56. E. J. Klekowski, H. G. Baker, Evolutionary significance of polyploidy in the Pteridophyta. *Science* **153**, 305–307 (1966).
57. C. H. Haufler, Homospory 2002: An Odyssey of Progress in Pteridophyte Genetics and Evolutionary Biology: Ferns and other homosporous vascular plants have highly polyploid chromosome numbers, but they express traits following diploid models and although capable of extreme inbreeding, are predominantly outcrossing. *BioScience* **52**, 1081–1093 (2002).
58. S. Koren *et al.*, Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
59. C. L. Xiao *et al.*, MECAT: Fast mapping, error correction, and de novo assembly for single-molecule sequencing reads. *Nat. Methods* **14**, 1072–1074 (2017).
60. O. Dudchenko *et al.*, The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. bioRxiv [Preprint] (2018). https://doi.org/10.1101/254797 (Accessed 18 January 2018).
61. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
62. S. Mirarab *et al.*, ASTRAL: Genome-scale coalescent-based species tree estimation. *Bioinformatics* **30**, i541–i548 (2014).
63. H. Tang *et al.*, Synteny and collinearity in plant genomes. *Science* **320**, 486–488 (2008).
64. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of next generation sequencing data. *BMC Bioinform.* **15**, 356 (2014).
65. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
66. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
67. A. Zwaenepoel, Y. Van de Peer, wgd-simple command line tools for the analysis of ancient whole-genome duplications. *Bioinformatics* **35**, 2153–2155 (2019).
68. Z. Li *et al.*, Multiple large-scale gene and genome duplications during the evolution of hexapods. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4713–4718 (2018).
69. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
70. C. Li *et al.*, de novo genome assembly of *Huperzia asiatica*. NGDC BioProject. https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA013778. Deposited 9 December 2022.
71. C. Li *et al.*, *Huperzia asiatica* genome. figshare. https://figshare.com/projects/Huperzia_asiatica_genome/169145. Deposited 6 June 2023.
72. JGI *et al.*, *Diphasiastrum complanatum* cultivar:PW_Plant_1. NCBI BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA914350. Deposited 8 March 2023.
73. JGI *et al.*, *Diphasiastrum complanatum* v3.1. phytozome. https://phytozome-next.jgi.doe.gov/info/Dcomplanatum_v3_1. Deposited 8 March 2023.