Evolution of genome space occupation in ferns: linking genome diversity and species richness

Tao Fujiwara[1,2], Hongmei Liu[3], Esteban I. Meza-Torres[4], Rita E. Morero[5], Alvaro J. Vega[4], Zhenlong Liang[1], Atsushi Ebihara[6], Ilia J. Leitch[7] and Harald Schneider[1*]

[1]Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Yunnan, China; [2]Makino Herbarium, Tokyo Metropolitan University, 1-1 Minami-osawa, Hachioji, Tokyo 192-0397, Japan.

[3]Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Yunnan, China; [4]Instituto de Botánica del Nordeste, Universidad Nacional del Nordeste, Consejo Nacional de Investigaciones Científicas y Técnicas, Corrientes, Argentina; [5]Instituto Multidiscipinario de Biologia Vegetal, Universidad Nacional de Cordoba, Consejo Nacional de Investigaciones Científicas y Tecnicas, Cordoba, Argentina; [6]Department of Botany, National Museum of Nature and Sciences, Tsukuba, Japan; [7]Jodrell Laboratory, Royal Botanic Gardens, Kew, London, UK.

**For correspondence. E-mail harald@xtbg.ac.cn*

- **Background and Aims** The dynamics of genome evolution caused by whole genome duplications and other processes are hypothesized to shape the diversification of plants and thus contribute to the astonishing variation in species richness among the main lineages of land plants. Ferns, the second most species rich lineages of land plants are highly suitable to test this hypothesis because of several unique features that distinguish fern genomes from those of seed plants. In this study, we tested the hypothesis that genome diversity and disparity shape fern species diversity by recording several parameters related to genome size and chromosome number.
- **Methods** We conducted *de novo* measurement of DNA C-values across the fern phylogeny to reconstruct the phylogenetic history of the genome space occupation in ferns by integrating genomic parameters such as genome size, chromosome number, and average DNA amount per chromosome into a time-scaled phylogenetic framework. Using phylogenetic generalized least square methods, we determined correlations between chromosome number and genome size, species diversity and evolutionary rates of their transformation.
- **Key Results** The measurements of DNA C-values for 233 species more than doubled the taxon coverage from ca. 2.2% in previous studies to 5.3% of extant diversity. The dataset documented not only substantial differences in the accumulation of genomic diversity and disparity among the major lineages of ferns but also recovered support the predicted correlation between species diversity and the dynamics of genome evolution.
- **Conclusions** Our results demonstrated substantial genome disparity among different groups in ferns and supported the prediction that alterations of reproductive modes alter trends of genome evolution. Finally, we recovered evidence for a close link between the dynamics of genome evolution and species diversity in ferns for the first time.

Key words: DNA C-values, genome evolution, genome size, macroevolution, polyploidy, pteridophytes.

INTRODUCTION

Increasing evidence demonstrates a close correlation between the evolvability—the innate capacity of a lineage to evolve by adapting to changes in environmental conditions (Kirchner and Gerhart 1998; Pigliucci 2008; Rabosky *et al.* 2013)—and the dynamics of genome evolution (Leitch and Leitch 2012; Bromham *et al.* 2015; Puttick *et al.* 2015) with the latter shaped by processes such as whole genome duplications [WGD], amplification of repetitive DNA [RDA], and diploidization involving DNA deletions [DPD] (Freeling *et al.* 2015; Soltis *et al.* 2015; Schubert and Vu 2016). These processes have been considered not only to cause the "C-Value Enigma" —a 64,000-fold genome size disparity in eukaryotes and 2,400-fold disparity in plants—but also to contribute to the remarkable differences in species richness among major lineages of eukaryotes and land plants (Gregory 2004; Leitch and Leitch 2013; Slijepcevic 2018; Choi *et al.* 2020). Indeed, these arguments have been discussed as a consequence of observed differences among land plant lineages in their genome diversity as indicated by holoploid genome size, chromosome numbers, and composition of genomic components such as non-coding DNA elements (Leitch and Leitch 2012; Pellicer *et al.* 2018). Whereas studies focused so far on angiosperms and gymnosperms have recovered substantial differences in genome diversity, less attention has been given to ferns, despite their genomes characterized by several unique features.

Containing ca. 11,000 extant species (PPGI 2016), ferns are not only the second most species-rich lineages of vascular plants, but are characterized by (1) the highest frequency of polyploidy enforced speciation events (Wood *et al.* 2009), (2) accumulation of large chromosome numbers including the largest number of chromosomes recorded among all organisms (Khandelwal 1990; Clark *et al.* 2016), (3) accumulation of medium to large genome sizes including one of the largest genomes recorded (Hidalgo *et al.* 2017; Pellicer *et al.* 2018), (4) a positive correlation of genome size with chromosome number and LTR-RT insertion time, respectively (Nakazato *et al.* 2008; Barker 2013; Baniaga and Barker 2019), and (5) a high rate of recurrent WGD during the phylogenetic history of several fern lineages (Clark *et al.* 2016; Huang *et al.* 2020). Lastly, the first complete and partially sequenced fern genomes recovered evidence for recurrent WGD and spread of

repeat elements in homosporous ferns (Wolf *et al.* 2015; Marchant *et al.* 2019), whereas the heterosporous ferns were distinct in their relatively small genome size and composition of non-coding DNA (Li *et al.* 2018). These results were consistent with the predictions made by previous investigations using different evidence to explore fern genomes (Barker and Wolf 2010; Barker 2013; Haufler 2014; Clark *et al.* 2016). Due to the unique characteristics observed in fern genomes and the phylogenetic position as sister lineage of seed plants, an exploration of fern genome space has been expected to shed new light on genome evolution in land plants (Rensing 2017). Besides, ferns show not only the diversity of life histories (Petersen and Burd 2017), adaptation to low-light condition (Schneider *et al.* 2004), and ecological adaptive strategies (Schuettpelz and Pryer 2009), but also a highly uneven phylogenetic distribution of extant species diversity in which 78% of the extant fern diversity belonging to only three lineages in Polypodiales (PPGI 2016). Together, these characteristics set ferns to be highly suitable to test the association between genome diversity and lineage diversification.

In this study, we reconstructed evolutionary patterns of genome size across the whole fern phylogeny with the aim to explore the hypothesis that genome diversity and disparity shape fern species diversity. Specifically, the experimental setup was designed to expand the phylogenetic and taxonomic coverage of DNA C-values as required to explore the fern genome space occupation using five parameters—holoploid genome size (1C), gametic chromosome number (n), monoploid genome size (1C*x*), basic chromosome number (*x*), and average DNA amount per chromosome (1C/n). The latter value enabled to quantify the packaging of DNA in form of chromosomes. In general, the fern genome has been considered to be very stable based on the observation of a positive correlation between chromosome number and genome size, which was interpreted as a consequence of recurrent WGD combined with conservation of chromosome structure (CCS) and delayed diploidization (Nakazato *et al.* 2008; Barker 2013; Clark *et al.* 2016). This correlation was confirmed using the highly improved dataset. In turn, we determined the phylogenetic pattern of observed genome size variation because lineage-specific traits and historical events were expected to contribute to the

accumulated genome disparity of ferns. This prediction is supported by the observations on genome size variation among lineages of ferns (Dyer *et al.* 2013; Clark *et al.* 2016; Liu *et al.* 2019) and the reported differences of genomes of heterosporous ferns from those in their homosporous relatives (Klekowski and Baker 1966; Haufler 2014; Li *et al*. 2018). Finally, the hypothesis that the genome diversity shapes fern species diversity was tested by specifically focusing on a link between species richness and the dynamics of genome evolution estimated as a rate of genome size evolution.

MATERIALS AND METHODS

Taxon sampling

DNA C-values were obtained with the aim to double at least the number of species with reported genome sizes and to cover particularly lineages lacking data until now. The latest fern classification PPGI (2016) was employed as guide to obtain species estimates and a robust phylogenetic framework. New genome size measurements were generated from materials collected mostly in southern China and southern South America by combining fieldwork and utilization of living collections at the Xishuangbanna Tropical Botanical Garden (China) and Wuhan Botanical Garden (China), and the Tsukuba Botanical Garden (Japan). These expanded sampling enabled us to overcome the bias towards species occurring in temperate regions of Europe and North America that hampered previous analyses (e.g., Clark *et al.* 2016).

De novo *measurement of genome size and chromosome counts*

DNA C-values were obtained using flow cytometry with propidium iodide following preparation procedures as described previously (Clark *et al.* 2016). Leaf fragments of the studied taxon were chopped together with an internal standard—either *Glycine max* 'Polanka' (2C = 2.50 pg) (Doležel *et al.* 1994), *Pisum sativum* 'Ctirad' (2C = 9.09 pg) (Doležel *et al.* 1992), or *Vicia faba* 'Inovec' (2C = 26.90 pg) (Doležel *et al.* 1992)—in an isolation buffer—either General Purpose Buffer

(Loureiro *et al.* 2006) or Ebihara Buffer (Ebihara *et al.* 2005). The fluorescence intensities were analyzed on BD FACSVerseTm (BD Bioscience, San Jose, CA, USA). Each taxon was measured at least three times using leaves from same individual, if possible, to enable the calculation of mean 2C-values and standard deviations for each taxon based on the fluorescence ratios between the sample and internal standards. Genome size data obtained in this study is provided in Supplementary Data Table S1. New chromosome counts were carried out using the established protocol (Takamiya 1993; Supplementary Data Table S2 and Fig. S1-4).

Data extraction from databases and literatures

Besides the newly generated data, previously published DNA C-values were obtained via the "Pteridophyte DNA C-values database" (http://data.kew.org/cvalues/) (Bennett and Leitch 2005) and recent publications (Clark *et al.* 2016; Dauphin *et al.* 2016; Yahaya *et al.* 2016; Fujiwara *et al.* 2017, 2018; Chang *et al.* 2018; Li *et al.* 2018; Kuo and Li 2019; Supplementary Data Table S1). Chromosome counts were assembled via "Index to Plant Chromosome Number" (IPCN; http://www.tropicos.org/project/ipcn) (Goldblatt and Johnson 1979) and the "Chromosome Count Database"(CCDB; http://ccdb.tau.ac.il/; Rice *et al.* 2015).

Characterizing fern genomes

We inspected five genomic parameters: (1) holoploid genome size [1C], (2) gametic chromosome number [n], (3) basic chromosome number [*x*], (4) monoploid genome size [1C*x*], (5) average DNA amount per chromosome [1C/n]. Basic chromosome number (*x*) was determined as the lowest gametic chromosome number in each genus (Manton and Vida 1968), whilst ploidy level for each taxon was estimated based on the basic chromosome number for the genus. Most fern genera show highly conserved chromosome numbers but some genera exhibit extensive chromosome number variation, e.g. *Hymenophyllum* (Hennequin *et al.* 2010) and *Lepisorus* (Wang *et al.* 2010). In such a

case, we first inferred ploidy level for each taxon by comparing the chromosome number of the taxon with gametic chromosome number known in the group that the taxon belongs to. Subsequently, we obtained basic chromosome number for the taxon by dividing its chromosome number by its ploidy level. Monoploid genome size (1C*x*) is defined as genome size per basic chromosome number (Greilhuber *et al.* 2005) and was obtained by dividing 2C-value by ploidy level for each taxon. Average DNA amount per chromosome (1C/n) was calculated by dividing 1C by the gametic chromosome number (n).

Phylogenetic analysis

Total genomic DNA was extracted from silica gel-dried samples using EasyPure® Plant Genomic DNA Kit (Transgen Biotech, Beijing, China). We amplified chloroplast gene *rbcL* with the standard PCR protocols (Schuettpelz and Pryer 2007). Sanger sequencing was outsourced in BGI sequencing service (http://www.genomics.cn/en). We also collected *rbcL* and *atpB* sequences from GenBank. Out of 561 species with genome size data, 430 species were selected based on the availability of DNA sequences and all genomic parameters examined in this study, for the phylogenetic analysis. All sequences including *Selaginella helvetica* (AB574644), *Isoetes sinensis* (AB574660), and *Lycopodium clavatum* (AB574626) as outgroup, were merged into single sequence matrix and aligned using MAFFT (Katoh and Standley 2013), followed by manual editing in AliView (Larsson 2014). The best model of nucleotide substitution was selected with jModelTest 2 (Darriba *et al.* 2012). Tree reconstruction was performed using Maximum Likelihood (ML) as implemented in RAxML-HPC2 8.2.6 (Stamatakis 2014) on the CIPRES Science Gateway portal (http://www.phylo.org/). We used GTRGAMMA as substitution model and performed 1000 bootstrap (BS) replicates. Tree topology for phylogenetic location of orders and families was constrained based on PPGI (2016). The obtained ML tree is shown in Supplementary Data Fig. S5. To obtain ultrametric tree, we carried out the penalized likelihood method using PATHd8 (Smith and O'Meara 2012). We set 17 fixed fossils and second calibration points in the best maximum likelihood tree in RAxML

based on fossil records and the estimated ages from large scale fossil integrating study reconstructing the divergence times of ferns (Testo and Sundue 2016) (Supplementary Data Table S3). The obtained ultrametric tree is shown in Supplementary Data Fig. S6.

Statistical analysis

Each genomic parameter was explored with Shapiro-Wilk normality tests and qqplot in R package. Because all of them significantly deviated from normal distribution ($p < 0.001$), they were log transformed before subsequent analyses.

For phylogenetic signal for each genomic parameter, Pagel's λ was estimated using the fitContinuous function in the GEIGER package (Harmon *et al.* 2008) in R. Significance of estimated value was tested by likelihood ratio test (LRT) comparing with the value in lambda = 0 model. To detect the heterogeneity of genome size evolution, three a priori scenarios were tested under Brownian Motion (BM) and Ornstein-Uhlenbeck (OU) Process using OUwie (Beaulieu *et al.* 2012). The employed scenarios were organised from simple to increasing complex models: 1) scenario in which all ferns evolve under a same trend (Singular Model); 2) scenario in which heterosporous fern evolves under distinct trend from homosporous fern (Heterosporous Fern Model); and 3) scenario in which all orders evolve independently under distinct trends from each other (All Orders Model). The BM process is a random walk process with a pure stochastic change to any value in a trait space, regulated by only rate parameter (σ^2), while the OU process controls changes of value by incorporation of an attractor (α) that is the strength to move back to an optimum value (θ), together with σ^2. For OU process with multiple regimes, we applied four OU processes, OUM (only θ varies among distinct regimes), OUMV (θ and σ^2), OUMA (θ and α) and OUMVA (θ, α and σ^2) with different assumptions about what of three parameters, σ^2, α and θ vary among distinct regimes. For these analyses, we prepared multi-regimes phylogeny using paintSubTree function in the phytools package (Revell 2012). We conducted model selection for three genomic parameters—1C, 1C*x* and 1C/n—and selected the best model by calculation of a sample size corrected Akaike Information Criterion

(AICc). To discover additional rate shifts for 1C, 1C*x* and 1C/n, we conducted l1ou (Khabbazian *et al.* 2016) analysis that is a computationally efficient approach that uses a lasso method to automatically detect evolutionary shifts under OU process. We used the estimate_shift_configuration function in this package with pBIC model as selection criterion. We calculated bootstrap value for each shift location with 500 iterations using l1ou_bootstrap_support function. Same analyses were also conducted for reduced dataset containing only Polypodiales.

We examined the relationships between genomic parameters, (1C vs n, 1C vs 1C/n) by employing phylogenetic generalized least square (PGLS) analysis (Grafen 1989), as implemented by the pgls function with lambda = 'ML' in the caper package (Orme 2013), against four different groups, "All ferns", "Homosporous ferns", "Leptosporagiate ferns" and "Polypodiales". The same analytical settings were employed to examine the correlation between 1C*x* and *x* as necessary to detect a putative bias created by neo-polyploid taxa.

To examine the predicted correlations between total species number and evolutionary rates of genome size, PGLS analyses were performed under the same setting as described above. We basically conducted order-level comparison to examine it. However, because species diversity in ferns is highly biased towards Polypodiales that accounts almost 80 % of whole extant species diversity in ferns (PPGI 2016). To avoid the expected bias, the order Polypodiales was separated into families instead of a single unit. This treatment takes into account that species diversity is highly variable even within Polypodiales and its pattern of genome size evolution is also heterogeneous although other basal orders show more stable patterns (see Result and Discussion below). The evolutionary rates of 1C, 1C*x* and 1C/n for each order except for Polypodiales, and each family including genome size data for more than 3 species in the Polypodiales were estimated as σ^2 under Brownian motion calculated with fitContinuous function in the GEIGER package (Supplementary Data Table S4). In parallel, we also examined correlations between diversification rate and evolutionary rates of genome size. We adopted a conservative way to estimate diversification rate in each clade by calculating the rate under a single constant-rate model based only on clade age and species richness (see Magallón and Sanderson 2001). The diversification rates for each order, and each family of Polypodiales were

calculated using bd.ms function in the GEIGER package with the parameters, epsilon and missing set to zero. We used species richness and clade age for each group from PPGI (2016) and Testo and Sundue (2016), respectively. All values were log-transformed before analysis.

RESULTS

Genome size variation throughout ferns

Our *de novo* measurements from 233 species combined with previously published DNA C-values highly enhanced the taxon coverage from ca. 2.2% to ca. 5.3% including 100% of the orders and 50% of the fern genera (Table 1 and Supplementary Data Table S1). The taxon coverage varied across the fern phylogeny ranging from the lowest value of 2.8% in Hymenophyllales to the highest value of 80% in Equisetales (Table 1). Ferns showed a mean value of 12,377 Mb -in holoploid genome size (1C) ranging 629.5-fold from 234 Mb of *Salvinia cucullata* to 147,391 Mb of *Tmesipteris obliqua*, a mean of 8,599 and 387.6-fold range in monoploid genome size (1C*x*) and a mean of 223 and 108.7-fold range in average DNA amount per chromosome (1C/n) variation.

Genome size evolution throughout phylogeny of ferns

All genomic parameters examined showed significant phylogenetic signals ($p < 0.001$), specifically 1C: $\lambda = 0.907$, 1C*x*: $\lambda = 0.954$, 1C/n: $\lambda = 0.948$, n: $\lambda = 0.797$, x: $\lambda = 1.00$ (Supplementary Data Table S5). The OUwie selected All-order model with OU processes that was found to be better fit models for the three genomic parameters (1C, 1C*x* and 1C/n) than the other models (Table 2). For each of three genomic parameters, although OUMA and OUMVA model showed extremely higher log likelihood and lower AICc than any other models (Supplementary Data Table S6), the parameters values estimated in the models showed large deviations from those in other models and thus these models were discarded due to their inappropriate model fitting. Therefore, among OU processes, OUM that infers only different optimum values among different groups was selected as the best

model for 1C value and OUMV that infer a different rate parameter and optimum values for 1C*x* and 1C/n values (Table 2 and Supplementary Data Table. S6). The l1ou algorithm recovered several evolutionary shifts in genome size that occurred throughout the phylogeny for the three genomic parameters, showing high heterogeneity of genome evolution in ferns (Fig. 1 and Supplementary Data Fig. S7).

Correlation among genome size and chromosome number

Holoploid genome size (1C) was significantly positively correlated with gametic chromosome number (n) and average DNA amount per chromosome (1C/n) across the phylogeny of ferns including the major groups tested independently (Fig. 2A, Supplementary Data table S7 and Fig. S8). In each group tested, the chromosome number fitted better with the holoploid genome size than the 1C/n value as indicated by the λ-values > 0.900 and R^2 > 0.45. Although exclusion of the bias created by neo-polyploidy, monoploid genome size (1C*x*) was significantly correlated with basic chromosome number for all ferns and leptosporangiate ferns. However, this correlation was not supported for homosoporous ferns and Polypodiales (Fig. 2B and Supplementary Data table S7).

The rate of genome size evolution is correlated with species diversity

Significant positive correlations were observed between the total number of species and each evolutionary rate of the three genomic parameters: 1C rate, 1C*x* rate and 1C/n rate (Fig. 3A and Supplementary Data Table S8). Similar positive correlations were also identified between diversification rates and rates of genome size evolution (Fig. 3B). In contrast, there is no significant correlation between the total number of species and the mean values for the three genomic parameters (Supplementary Data Fig. S9 and Table S9).

DISCUSSION

The newly generated DNA C-values of 233 species highly improved taxon coverage to 5.3% including 100% of the orders and 50% of the fern genera according to PPGI (2016). This coverage is more than doubled compared with previous summaries, 2.3 % and 2.2 % in Clark et al. (2016) and Pellicer et al. (2018), respectively, and is the best among land plant lineages containing more than 2,000 species (Pellicer *et al.* 2018). Using our updated dataset, we examine genome size disparity throughout fern phylogeny and its association with species diversification.

Genome size disparity across the phylogeny of ferns

The three genomic parameters recruited to explore the genome space evolution—1C, 1C*x*, and 1C/n—show significantly high phylogenetic signals throughout fern phylogeny (Supplementary Data Table S5), suggesting genome space disparity among the main lineages. Consistent with this, OUwie analyses selected the All-order model with OU process, where each order has a distinct evolutionary trend from each other, as best models for all three genomic parameters, specifically OUM model for 1C, and OUMV model for 1Cx and 1C/n (Table 2 and Supplementary Data Table. S6). Because the OU process has been the better fit to all genomic parameters for genome size than the BM process, we concluded that each fern order evolved distinct optimal values of genome size and structure (Table .2). The selection of OUMV mode—assuming not only optimal value but also rate parameter as variable among distinct groups—as the best models for 1C*x* and 1C/n, suggests that different fern orders display differences in dynamics of monoploid genome size and chromosome size evolution. The ancient sister lineages, Ophioglossales and Psilotales share the accumulation of large genomes in 1C but show opposite trends in 1C/n (Table 1), where they converge to contrasting rates of genome size evolution and optimal values (694 Mb for Psilotales and 177 Mb for Ophioglossales, Table 2). Among homosporous ferns, Gleicheniales is distinct by its relatively small genomes that are the consequence of rate shift towards genome size reduction (Table 1, Fig. 1 and Table .2), and its 1C/n values is comparable to heterosporous ferns. On the other hand, other ancient lineages of

leptosporangiate ferns, namely Osmundales and Hymenophyllales evolved towards one of the largest optimal values of 1C/n among all ferns (721 Mb for Osmundales and 501 Mb for Hymenophyllales, Table 2), consequently accumulating large 1C and 1C*x* (Table 1). However, the two lineages show contrasting rates of genome size evolution in 1C*x* and 1C/n. Whereas Osmundales exhibits the second slowest rate of genome size evolution (see Schneider *et al.* 2015), Hymenophyllales experienced relatively faster genome change as consistent with the report of basic chromosome number variations (Hennequin *et al.* 2010). The largest optimal values of 1C/n were consistent with previous reports documenting very large chromosomes in at least some of these ferns (Lovis 1978). In contrast to other orders of ferns, the Polypodiales exhibits heterogenous trends of genome size evolution as a result of rate shift in genome size evolution within this order. Almost half of evolutionary shifts in genome size are located inside the order Polypodiales (5/10 shifts in 1C, 10/21 in 1C*x* and 9/18 in 1C/n see Fig. 1 and Supplementary Data Fig. S7). The frequency of shifts is consistent with the genome disparity recovered by the largest ranges of 1C and the second largest of 1C*x* and 1C/n among all ferns (Table 1). Most notable, the three most species rich fern lineages (PPGI 2016), namely Aspleniineae, Polypodiineae, and Pteridineae, contributing together ca. 78% of the extant fern diversity, are highly variable in the three genomic parameters (1C, 1C*x*, and 1C/n) recruited to elucidate the genome disparity).

Genome size disparity is not restricted to homosporous ferns but also occurs between homosporous and heterosporous ferns. Heterosporous ferns are distinct in their trend towards the smallest values in all three core parameters (1C, 1C*x*, and 1C/n) and the largest disparity and evolutionary rate of 1C*x* and 1C/n detected among ferns (Table 1). This result is consistent with the conclusions taken from the whole genome sequences of these ferns (Li et al. 2018). The rapid transformation of the genomes of these ferns is likely caused by the transition from homosporous to heterosporous reproduction. Thus, the observed result supports the hypothesis that changes in the reproductive system accelerated the rate of genome evolution in this relatively young lineage of ferns (Haufler 2014). However, it should be carefully noted that these results in 1C*x* and 1C/n may be misled as a consequence of misassignment of chromosome numbers and ploidy level to species in this

group. For example, although the reported genome sizes of *Salvinia cuculata* and *S. molesta* are 0.48 and 4.45 pg/2C respectively, the chromosome numbers and ploidy levels for both species were reported to be 2n = 45 and pentaploid level (Tatuno and Takei 1969). Thus, some of Salviniales species may show intraspecific chromosome variation that may cause the ambiguous link between genome size and chromosome number. Therefore, the link between genome size and chromosome number in Salviniales is required for further exploration to elucidate genome size disparity between homosporous and heterosporous ferns.

Repeated WGD with delayed DPD and CCS

Holoploid genome size (1C) was significantly positively correlated with both chromosome number (n) and average DNA amount per chromosome (1C/n) across the phylogeny of ferns including the major groups tested independently (Fig. 2A, Supplementary Data table S7 and Fig. S8). The model with chromosome number fitted better with 1C than that with 1C/n for all groups examined (Supplementary Data table S7). This result was consistent with previous reports that ferns were one of two lineages of land plants showing this correlation (Nakazato *et al.* 2008; Barker 2013; Clark *et al.* 2016). This pattern may be explained by repeated WGD in the phylogeny of ferns combined with delayed genome size reduction (such as DPD) as reflected by the conservation of chromosome structure (CCS). The later was elucidated using the 1C/n value. This conclusion was also supported by the observation of a conservation of LTR since their insertion (Baniaga and Barker 2019) and the spread of repeat elements in homosporous fern genomes (Marchant *et al.* 2019).

1C*x* was significantly positively correlated with basic chromosome number (*x*) across all ferns ($p < 0.01$) and leptosporangiate ferns ($p < 0.05$) but this correlation was weak as indicated by the R^2 ($R^2 = 0.015$ for all ferns and $R^2 = 0.014$) and rejected for homosporous ferns ($p = 0.162$) and Polypodiales ($p = 0.230$) (Fig. 2B and Supplementary Data table S7). This result was consistent with recently published results that the 1C/n was not as constant as some authors assumed despite a general trend towards the conservation for relatively small chromosomes in most fern lineages compared to

angiosperms and gymnosperms (Dyer *et al.* 2013; H-M Liu *et al.* 2019). Due to the high frequency of neopolyploidy combined with the trend to conserve chromosome structure, the contribution of processes such as DNA deletion and selective DNA amplification was elucidated only by focusing on monoploid genome size. Nonetheless, our analysis against a reduced dataset containing only Polypodiales found additional evidence supporting the breakdown of chromosome structure conservation (CCS) in this species richest fern lineage (Supplementary Data Fig. S10), where it experienced several shifts of the 1C/n value which largely contributed to contraction and expansion of the genome size. This conclusion was supported by the relatively small numbers of reported dysploidy sequences in ferns (Lovis 1978; Hennequin *et al.* 2010; Wang *et al.* 2010). These results suggest that deviation of the CCS was more common in species-rich lineages compared to the overall patterns recovered in ferns.

Genome size diversity and species richness

Total number of species and diversification rate were correlated with the evolutionary rate of three genomic parameters—1C, 1C*x* and 1C/n (Fig. 3 and Supplementary Data Table S8), but not with the mean values of them (Supplementary Data Fig. S9 and Table S9). This result resembled the reported correlation between the dynamics of genome evolution and diversification rates in angiosperms (Puttick *et al.* 2015). In ferns, the correlation may be mainly explained by the high frequency of polyploidy (Wood *et al.* 2009) because lineages showing high rates of 1C were known to include many polyploid species such as Ophioglossales (Khandelwal 1990; Dauphin *et al.* 2018), Hymenophyllales (Ebihara *et al.* 2005; Nitta *et al.* 2011), and the most of families in Polypodiales, e.g. Aspleniaceae, (Schneider *et al.* (2017). However, polyploidy did not explain the observed patterns alone, because we also found the correlation between species richness and the evolutionary rate of 1C*x* and 1C/n. These findings support that other processes such as DPD and RDA may have contributed to fern diversification but their impact has been substantially lower compared to angiosperms (Schubert and Vu 2016). In particular, the species rich lineages of Polypodiales showed

evidence for enhanced rates of genome evolution as a consequence of equal contribution of WGD, RDA, and DPD with the consequence of enhanced diversification rates that contributed to the evolutionary success of this fern lineage. This result supports the recurrent WGD during the diversification of derived ferns as suggested based on transcriptome data (One Thousand Plant Transcriptomes Initiative 2019; Huang *et al.* 2020). In turn, the less dynamic genome space exploration may also explain the rather small species diversity of ferns compared to angiosperms.

CONCLUSION

In summary, our results provide strong support to the hypothesis that the diversification of plant lineages has been shaped by the dynamics of their genome evolution. Instead of focusing on whole genome sequences or transcriptomes, the parameters linked to the amount of DNA and its packaging in chromosomes were used to elucidate the dynamics of genome space evolution in ferns. This approach has the disadvantage that we cannot trace changes in genome composition such as the contribution of different kinds of repetitive DNA or the fate of duplicated genes originated in ancient whole genome duplication to genome evolution. However, this disadvantage has to be taken in the context of the much denser and more balanced taxon sampling achieved than in any study using sequenced genomes alone (Li *et al.* 2018; One Thousand Plant Transcriptomes Initiative 2019). Future studies will hopefully integrate genomic and transcriptomic sequence evidence into the framework created using the parameters used in this study. Despite the limited genomic sequence data available for ferns, our results are highly consistent with studies that used complete or partial sequenced genomes of heterosporous and homosporous ferns (Wolf *et al*.2015; Li *et al*. 2018). Furthermore, we confirmed several predictions based on the hypothesis that different trends in chromosome and genome evolution between homosporous and heterosporous ferns could be attributed to their reproductive system, proposed in the early studies on fern genetics (Klekowski and Baker 1966; Haufler 2014).

Expanding our findings, the remarkable difference in species richness among land plant may be explained by differences in the dynamics of genome evolution. The remarkable success of angiosperms exceeding all other land plant lineages in their species richness, phenotypic diversity, and ecological importance is arguably the consequence of innovations enabling much faster rates of genome evolution compared to their sister lineages, the gymnosperms and more distant relatives, ferns and bryophytes (Leitch and Leitch 2012; Puttick *et al.* 2015; Pellicer *et al.* 2018). In turn, lower rate of genome evolution may explain the lower species numbers of ferns. However, the enhanced rate of genome evolution also promoted the success of some recently diverging lineages. Whole genome duplications alone fail to explain these successes despite their prominence because other process, e.g. amplification and purge of repetitive DNA, and diploidization with DNA deletion, may had a crucial contribution to unleash polyploidy driven innovations.

ACKNOWLEDGEMENTS

Authors would like to thank Jaroslav Doležel (Institute of Experimental Botany, the Czech Academy of Sciences) for providing reference plant standards for flow cytometry. H.S. is the leader of this project. T.F., H.L. and H.S. designed the experiments. H.L., A.E. and I.L. provided access to cultivated materials; T.F., H.L., H.S. Z.L., E.M., R.M., and A.E. collected samples. T.F., Z.L., E.M., R.M., and A.V. generated DNA C-Values, chromosome counts, and DNA sequences. T.F. carried out all analyses. T.F., H.L., and H.S. drafted the manuscript with the contribution of all authors. All authors gave final approval for publication.

FUNDING

This work was supported by the postdoctoral fellowships granted to T.F. by the Chinese Academy of Sciences (2019PB0148) and the Yunnan Province (Y7YN021B13), and research funds provided to H. L. and H.S. by Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences (grant numbers 2017XTBG-T03, 2017XTBG-F05).

LITERATURE CITED

Baniaga AE, Barker MS. **2019**. Nuclear genome size is positively correlated with median LTR-RT insertion time in fern and lycophyte genomes. *American Fern Journal* **109**: 248–266.

Barker MS. **2013**. Karyotype and genome evolution in pteridophytes In: *Plant Genome Diversity Volume 2*. Springer, 245–253.

Barker MS, Wolf PG. **2010**. Unfurling fern biology in the genomics age. *BioScience* **60**: 177–185.

Beaulieu JM, Jhwueng DC, Boettiger C, O'Meara BC. **2012**. Modeling stabilizing selection: Expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* **66**: 2369–2383.

Bennett MD, Leitch IJ. **2005**. Plant genome size research: A field in focus. *Annals of Botany* **95**: 1–6.

Bromham L, Hua X, Lanfear R, Cowman PF. **2015**. Exploring the Relationships between mutation rates, life history, genome size, environment, and species richness in flowering plants. *The American Naturalist* **185**: 507–524.

Chang Y, Ebihara A, Lu S, Liu H, Schneider H. **2018**. Integrated taxonomy of the *Asplenium normale* complex (Aspleniaceae) in China and adjacent areas. *Journal of Plant Research* **131**: 573–587.

Choi IY, Kwon EC, Kim NS. **2020**. The C- and G-value paradox with polyploidy, repeatomes, introns, phenomes and cell economy. *Genes and Genomics* **42**: 699–714.

Clark J, Hidalgo O, Pellicer J, *et al.* 2016. Genome evolution of ferns : evidence for relative stasis of genome size across the fern phylogeny. *New Phytologist* **230**: 1072–1082.

Darriba D, Taboada GL, Doallo R, Posada D. **2012**. jModelTest 2: more models, new heuristics and high- performance computing. *Nature Methods* **9**: 772.

Dauphin B, Grant JR, Farrar DR, Rothfels CJ. **2018**. Rapid allopolyploid radiation of moonwort ferns (*Botrychium;* Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. *Molecular Phylogenetics and Evolution* **120**: 342–353.

Dauphin B, Grant J, Mráz P. **2016**. Ploidy level and genome size variation in the homosporous ferns *Botrychium* s.l. (Ophioglossaceae). *Plant Systematics and Evolution* **302**: 575–584.

Doležel J, Doleželová M, Novák FJ. **1994**. Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *M . balbisiana*). *Biologia Plantarum* **36**: 351–357.

Doležel J, Sgorbati S, Lucretti S. **1992**. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* **85**: 625–631.

Dyer RJ, Pellicer J, Savolainen V, Leitch IJ, Schneider H. **2013**. Genome size expansion and the relationship between nuclear DNA content and spore size in the *Asplenium monanthes* fern complex (Aspleniaceae). *BMC Plant Biology* **13**: 1–13.

Ebihara A, Ishikawa H, Matsumoto S, *et al.* 2005. Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. *American Journal of Botany* **92**: 1535–1547.

Freeling M, Scanlon MJ, Fowler JF. **2015**. Fractionation and subfunctionalization following genome duplications: Mechanisms that drive gene content and their consequences. *Current Opinion in Genetics and Development* **35**: 110–118.

Fujiwara T, Serizawa S, Watano Y. **2018**. Phylogenetic analysis reveals the origins of tetraploid and hexaploid species in the Japanese *Lepisorus thunbergianus* (Polypodiaceae) complex. *Journal of Plant Research* **131**: 945–959.

Fujiwara T, Uehara A, Iwashina T, *et al.* 2017. Allotetraploid cryptic species in *Asplenium normale* in the Japanese Archipelago, detected by chemotaxonomic and multi-locus genotype approaches. *American Journal of Botany* **104**: 1390–1406.

Goldblatt P, Johnson DE. **1979**. Index to plant chromosome numbers (IPCN). *Missouri Botanical Garden website: http://mobot. mobot. org/W3T/search/ipcn. html*.

Grafen A. **1989**. The phylogenetic regression. *Philosophical Transactions of the Royal Society B: Biological Sciences* **326**: 119–157.

Gregory TR. **2004**. Insertion-deletion biases and the evolution of genome size. *Gene* **324**: 15–34.

Greilhuber J, Doležel J, Lysák MA, Bennett MD. **2005**. The origin, evolution and proposed stabilization of the terms “genome size” and “C-value” to describe nuclear DNA contents. *Annals of Botany* **95**: 255–260.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. **2008**. GEIGER: Investigating evolutionary radiations. *Bioinformatics* **24**: 129–131.

Haufler CH. **2014**. Ever since Klekowski: Testing a set of radical hypotheses revives the genetics of ferns and lycophytes. *American Journal of Botany* **101**: 1–7.

Hennequin S, Ebihara A, Dubuisson JY, Schneider H. **2010**. Chromosome number evolution in *Hymenophyllum* (Hymenophyllaceae), with special reference to the subgenus *Hymenophyllum*. *Molecular Phylogenetics and Evolution* **55**: 47–59.

Hidalgo O, Pellicer J, Christenhusz MJM, Schneider H, Leitch IJ. **2017**. Genomic gigantism in the whisk-fern family (Psilotaceae): *Tmesipteris obliqua* challenges record holder Paris japonica. *Botanical Journal of the Linnean Society* **183**: 509–514.

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

Katoh K, Standley DM. **2013**. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Khabbazian M, Kriebel R, Rohe K, Ané C. **2016**. Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. *Methods in Ecology and Evolution* **7**: 811–824.

Khandelwal S. **1990**. Chromosome evolution in the genus *Ophioglossum* L. *Botanical Journal of the Linnean Society* **102**: 205–217.

Kirschner M, Gerhart J. 1998. Evolvability. *Proceedings National Academy of Sciences, USA* **95**: 8420–8427.

Klekowski EJ, Baker HG. **1966**. Evolutionary Significance of Polyploidy in the Pteridophyta. *Science* **153**: 305–308.

Kuo L-Y, Li F-W. **2019**. A roadmap for fern genome sequencing. *American Fern Journal* **109**: 212–223.

Larsson A. **2014**. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**: 3276–3278.

Leitch AR, Leitch IJ. **2012**. Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytologist* **194**: 629–646.

Leitch IJ, Leitch AR. **2013**. Genome size diversity and evolution in land plants In: *Plant Genome Diversity Volume 2*. Springer, 307–322.

Li FW, Brouwer P, Carretero-Paulet L, *et al.* 2018. Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants* **4**: 460–472.

Liu HM, Ekrt L, Koutecky P, *et al.* 2019. Polyploidy does not control all: Lineage-specific average chromosome length constrains genome size evolution in ferns. *Journal of Systematics and Evolution* **57**: 418–430.

Loureiro J, Rodriguez E, Dolezel J, Santos C. **2006**. Comparison of four nuclear isolation buffers for plant DNA flow cytometry. *Annals of Botany* **98**: 679–689.

Lovis JD. **1978**. Evolutionary patterns and processes in ferns. *Advances in Botanical Research* **4**: 229–415.

Magallón S, Sanderson MJ. **2001**. Absolute diversification rates in angiosperm clades. *Evolution* **55**: 1762-1780.

Manton I, Vida G. **1968**. Cytology of the fern flora of Tristan da Cunha. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **170**: 361–379.

Marchant DB, Sessa EB, Wolf PG, *et al.* 2019. The C-Fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. *Scientific Report*s **9**: 18181.

Nakazato T, Barker MS, Rieseberg LH, Gastony GJ. **2008**. Evolution of the nuclear genome of ferns and lycophytes In: *Biology and evolution of ferns and lycophytes*. Cambridge University Press, 175–198.

Nitta JH, Ebihara A, Ito M. **2011**. Reticulate evolution in the *Crepidomanes minutum* species complex (Hymenophyllaceae). *American journal of botany* **98**: 1782–800.

One Thousand Plant Transcriptomes Initiative. **2019**. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679–685.

Orme D. **2013**. The caper package: comparative analysis of phylogenetics and evolution in R. *R package version* **5**: 1–36.

Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. **2018**. Genome size diversity and its impact on the evolution of land plants. *Genes* **9**: 88.

Petersen KB, Burd M. **2017**. Why did heterospory evolve? *Biological Reviews* **92**: 1739–1754.

Pigliucci M. **2008**. Is evolvability evolvable? *Nature Reviews Genetics* **9**: 75–82.

PPGI. **2016**. A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution* **54**: 563–603.

Puttick MN, Clark J, Donoghue PCJ. **2015**. Size is not everything: Rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20152289.

Rabosky DL, Santini F, Eastman J, *et al.* 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications* **4**: 1958.

Rensing SA. **2017**. Why we need more non-seed plant models. *New Phytologist* **216**: 355–360.

Revell LJ. **2012**. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.

Rice A, Glick L, Abadi S, *et al.* 2015. The Chromosome Counts Database (CCDB) - a community resource of plant chromosome numbers. *New Phytologist* **206**: 19–26.

Schneider H, Liu HM, Chang YF, *et al.* 2017. Neo- and Paleopolyploidy contribute to the species diversity of *Asplenium*—the most species-rich genus of ferns. *Journal of Systematics and Evolution* **55**: 353–364.

Schneider H, Liu H, Clark J, *et al.* 2015. Are the genomes of royal ferns really frozen in time? Evidence for coinciding genome stability and limited evolvability in the royal ferns. *New Phytologist* **207**: 10–13.

Schneider H, Schuettpelz E, Pryer KM, Cranfill R, Magallón S, Lupia R. **2004**. Ferns diversified in the shadow of angiosperms. *Nature* **428**: 553–557.

Schubert I, Vu GTH. **2016**. Genome stability and evolution: Attempting a holistic view. *Trends in Plant Science* **21**: 749–757.

Schuettpelz E, Pryer KM. **2007**. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon* **56**: 1037–1050.

Schuettpelz E, Pryer KM. **2009**. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences* **106**: 11200–11205.

Slijepcevic P. **2018**. Genome dynamics over evolutionary time: "C-value enigma" in light of chromosome structure. *Mutation Research - Genetic Toxicology and Environmental Mutagenesis* **836**: 22–27.

Smith SA, O'Meara BC. **2012**. TreePL: Divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**: 2689–2690.

Soltis PS, Marchant DB, Van de Peer Y, Soltis DE. **2015**. Polyploidy and genome evolution in plants. *Current Opinion in Genetics and Development* **35**: 119–125.

Stamatakis A. **2014**. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Takamiya M. **1993**. Chromosome Numbers of *Woodsia kitadakensis* and *W. subcordata* (W oodsiaceae). *Journal of Japanese botany* **68**: 73–76.

Tatuno S, Takei M. **1969**. Cytological studies of Salviniaceae. I. Karyotype of two species in the genus Salvinia. *The Botanical Magazine, Tokyo* **82**: 403–408.

Testo W, Sundue M. **2016**. A 4000-species dataset provides new insight into the evolution of ferns A 4000-species dataset provides new insight into the evolution of ferns. *Molecular Phylogenetics and Evolution* **105**: 200–211.

Wang L, Qi X, Xiang Q, Heinrichs J, Schneider H, Zhang X. **2010**. Phylogeny of the paleotropical fern genus *Lepisorus* (Polypodiaceae, Polypodiopsida) inferred from four chloroplast DNA regions. *Molecular phylogenetics and evolution* **54**: 211–225.

Wolf PG, Sessa EB, Marchant DB, *et al.* 2015. An exploration into fern genome space. *Genome Biology and Evolution* **7**: 2533–2544.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. **2009**. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 13875–13879.

Yahaya NH, Stech M, Zonneveld BJM, Hovenkamp PH. **2016**. What is *Nephrolepis* "bostoniensis"?. Unravelling the origin of *Nephrolepis* hybrids and cultivars with molecular data. *Scientia Horticulturae* **204**: 153–160.

FIGURE LEGENDS

Fig. 1 Phylogeny with clades painted to reflect rate shifts under Ornstein–Uhlenbeck process suggested by 1lou method for holoploid genome size (1C) (A) and average DNA amount per chromosome (1C/n) (B). Black squares and numbers indicate the locations of rate shift and bootstrap supports for these shift placements. Red and blue arrow near black squares indicates up and down shift in the rate shift. Bar plot located next to each phylogeny depicts the parameter value for each species and different colors show different orders. Each picture on left side represents each order. Classification according to PPGI (2016).

Fig. 2 Scatter plot showing correlation between holoploid genome size (1C = y-axis) and gametic chromosome number (x = x-axis) (A) and correlation between monoploid genome size (1C*x* = y-axis) and basic chromosome number (x = x-axis) (B) for "All ferns", "Homosporous ferns", "Leptosporangiate ferns", "Polypodiales". Each dot signifies one taxon. Each color corresponds to each order according to the right legend. Dashed lines indicate the regression lines calculated using PGLS. Classification according to PPGI (2016).

Fig. 3 Scatter plot showing correlation between each of total number of species (A) and diversification rate (B), and evolutionary rate of three genomic parameters, holoploid genome size (1C) (left), the monoploid genome size (1C*x*) (middle), and average DNA amount per chromosome (1C/n) (right). Classification according to PPGI (2016). Solid and dashed lines indicate the regression lines calculated for 'All ferns' and 'Homosporous ferns' using PGLS.

Table 1. Summary of mean, minimum (Min), maximum (Max) size and size range of holoploid genome size (1C), monoploid genome size (1C*x*) and average DNA amount per chromosome (1C/n). Total species number is based on PPGI (2016).

Order	SN	SN - 1C	TC - 1C (%)	Mean 1C (Mb)	Min 1C (Mb)	Max 1C (Mb)	1C size range	Mean 1C*x* (Mb)	Min 1C*x* (Mb)	Max 1C*x* (Mb)	1C*x* size range	Mean 1C/n (Mb)	Min 1C/n (Mb)	Max 1C/n (Mb)	1C/*n* size range
All ferns	10578	561	5.3	12377	234	147291	629.45	8599	95	36823	387.61	223	10	1087	108.7
Homosporous ferns	10388	552	5.3	12501	1823	14729	80.79	8695	1823	36823	20.2	225	44	1087	24.7
Equisetales	15	12	80	22517	11066	31833	2.88	22516	11066	31833	2.88	208	102	294	2.88
Ophioglossales	112	23	20.5	23140	7878	64108	8.14	10299	4231	14455	3.42	220	73	321	4.4
Psilotales	17	3	17.6	87919	60117	147291	2.45	34752	30053	36823	1.23	668	577	708	1.23
Marattiales	111	12	10.8	9685	4430	20548	4.64	7262	4425	10748	2.43	183	113	269	2.38
Osmundales	18	9	50	15337	13158	20547	1.56	15337	13158	20547	1.56	697	598	933	1.56
Hymenophyllales	434	12	2.8	19857	10494	31408	2.99	16661	5247	25236	4.81	559	146	1088	7.45
Gleichniales	172	13	7.6	3978	1824	18934	10.38	2522	1824	3183	1.75	59	40	80	2
Schizaeles	190	11	5.8	12594	5555	22939	4.13	9120	4675	14484	3.1	314	199	483	2.43
Salviniales	82	9	11	1405	235	2572	10.94	1064	96	1917	19.97	75	11	202	18.36
Cyatheales	713	23	3.2	10257	2465	24054	9.76	8473	2465	13868	5.63	125	44	210	4.77
Polypodiales	8714	434	5	11185	2380	59164	24.86	7695	2308	32758	14.19	211	61	798	13.08

Saccolomatineae	18	1	5.6	37907	-	-	-	9477	-	-	-	205	-	-	-
Lindsaeineae	234	7	3	7464	3550	12959	3.65	4841	3156	7540	2.39	105	66	151	2.29
Pteridiineae	1211	66	5.5	9989	2582	34099	13.21	5912	2308	13435	5.82	197	80	448	5.6
Dennstaedtiineae	265	12	4.5	7755	3166	14802	4.68	4686	3166	8963	2.83	137	66	168	2.55
Aspleniineae	2775	147	5.3	9755	2380	26558	11.16	6229	2380	19746	8.3	170	61	581	9.52
Polypodiineae	4208	201	4.8	13209	3702	59164	15.98	10003	3707	32758	8.84	266	100	799	7.99

SN: species number, SN-1C: number of species with reported 1C genome size, TC-1C: taxon coverage in 1C genome size

Table 2. Parameter estimates in the best models for 1C, 1C*x* and 1C/n from OUwie analysis. Optimum value (**θ**) in each genomic parameter was back-transformed from logistic value.

Genomic parameter	**1C**			**1C*x***			**1C/n**		
Best model	**OUM**			**OUMV**			**OUMV**		
Parameters	**α**	**σ^2**	**θ (Mb)**	**α**	**σ^2**	**θ (Mb)**	**α**	**σ^2**	**θ (Mb)**
Equisetales	0.7869	0.7069	20039	0.7869	0.1592	20039	0.7869	0.1592	186
Psilotales	0.7869	0.7069	72127	0.7869	0.0003	36064	0.7869	0.0003	694
Ophioglossales	0.7869	0.7069	18639	0.7869	0.2938	8832	0.7869	0.4116	177
Marattiales	0.7869	0.7069	8157	0.7869	0.0489	7267	0.7869	0.0486	183
Osmundales	0.7869	0.7069	15862	0.7869	0.0241	15862	0.7869	0.0241	721
Hymenophyllales	0.7869	0.7069	19536	0.7869	0.2932	17079	0.7869	0.3836	501
Gleicheniales	0.7869	0.7069	2501	0.7869	0.0706	2501	0.7869	0.0774	56
Schizaeales	0.7869	0.7069	13198	0.7869	0.0625	9498	0.7869	0.1407	295
Salviniales	0.7869	0.7069	887	0.7869	1.05	711	0.7869	1.1006	51
Cyatheales	0.7869	0.7069	8240	0.7869	0.3673	7634	0.7869	0.2935	116
Polypodiales	0.7869	0.7069	10400	0.7869	0.415	7746	0.7869	0.4064	215

Figure 1

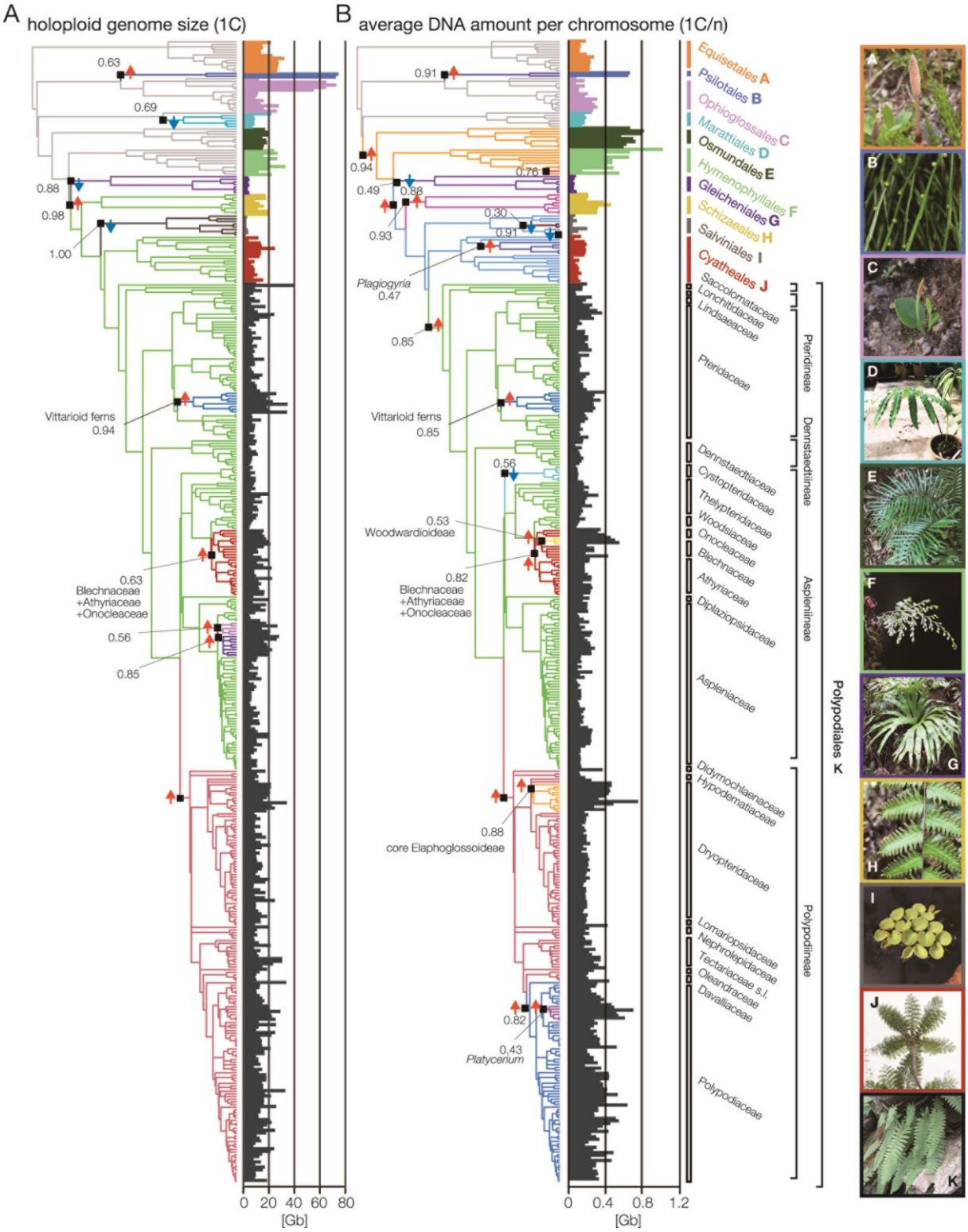

Figure 2

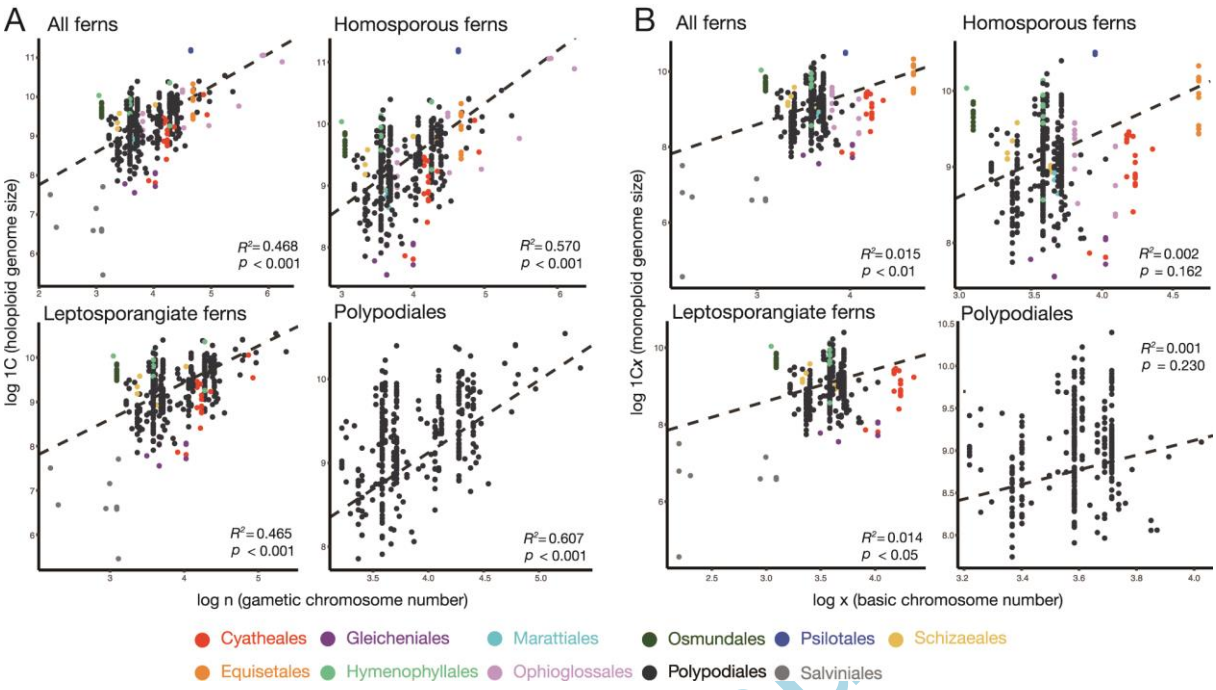

Figure 3