the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2024) **117,** 177–192 doi: 10.1111/tpj.16486

Genome sequences and population genomics provide insights into the demographic history, inbreeding, and mutation load of two 'living fossil' tree species of *Dipteronia*

Yu Feng[1,2,6], Hans Peter Comes[3], Jun Chen[1], Shanshan Zhu[1,4], Ruisen Lu[5], Xinyi Zhang[1,6], Pan Li[1], Jie Qiu[7], Kenneth M. Olsen[8,*] and Yingxiong Qiu[6,*]

[1]*Systematic & Evolutionary Botany and Biodiversity group, College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang, China,*

[2]*CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization & Ecological Restoration and Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China,*

[3]*Department of Environment & Biodiversity, Salzburg University, Salzburg, Austria,*

[4]*State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Ningbo University, Ningbo, Zhejiang 315211, China,*

[5]*Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing 210014, China,*

[6]*Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, Hubei 430074, China,*

[7]*Shanghai Key Laboratory of Plant Molecular Sciences, College of Life Sciences, Shanghai Normal University, Shanghai 200234, China, and*

[8]*Department of Biology, Washington University in St Louis, St Louis, Missouri 63130, USA*

Received 17 November 2022; revised 29 August 2023; accepted 20 September 2023; published online 5 October 2023.
*For correspondence (e-mail qyxhero@zju.edu.cn, qiuyingxiong@wbgcas.cn; kolsen@wustl.edu).

SUMMARY

'Living fossils', that is, ancient lineages of low taxonomic diversity, represent an exceptional evolutionary heritage, yet we know little about how demographic history and deleterious mutation load have affected their long-term survival and extinction risk. We performed whole-genome sequencing and population genomic analyses on *Dipteronia sinensis* and *D. dyeriana*, two East Asian Tertiary relict trees. We found large-scale genome reorganizations and identified species-specific genes under positive selection that are likely involved in adaptation. Our demographic analyses suggest that the wider-ranged *D. sinensis* repeatedly recovered from population bottlenecks over late Tertiary/Quaternary periods of adverse climate conditions, while the population size of the narrow-ranged *D. dyeriana* steadily decreased since the late Miocene, especially after the Last Glacial Maximum (LGM). We conclude that the efficient purging of deleterious mutations in *D. sinensis* facilitated its survival and repeated demographic recovery. By contrast, in *D. dyeriana*, increased genetic drift and reduced selection efficacy, due to recent severe population bottlenecks and a likely preponderance of vegetative propagation, resulted in fixation of strongly deleterious mutations, reduced fitness, and continuous population decline, with likely detrimental consequences for the species' future viability and adaptive potential. Overall, our findings highlight the significant impact of demographic history on levels of accumulation and purging of putatively deleterious mutations that likely determine the long-term survival and extinction risk of Tertiary relict trees.

Keywords: conservation genomics, demographic history, mutation load, population genomics, tertiary relict species.

INTRODUCTION

In most lineages, new species are formed while others become extinct, yet some ancient lineages of low taxonomic diversity, often termed 'living fossils', have neither gone extinct nor produced many new species and have remained relatively unchanged morphologically over long periods of evolutionary time (Lidgard & Love, 2018; Vargas et al., 2020; Werth & Shear, 2014). Among the northern temperate regions, East Asia harbors the highest richness of 'living fossil' tree genera (e.g., *Ginkgo*, *Metasequoia*,

Cercidiphyllum, and *Dipteronia*); yet, in contrast to their generally wider distributions in the Northern Hemisphere during the Tertiary, extant populations of these Tertiary relict trees are often small and isolated (Milne & Abbott, 2002). This implies an increased risk of extinction due to impacts on their genetic makeup from human activities, climate change, and/or demographic stochasticity (Charlesworth & Willis, 2009; Kardos et al., 2018; Lynch et al., 1995; Norén et al., 2016). Of the few genomic studies that have focused on East Asian living fossil trees to date (e.g., *Ginkgo biloba*: Zhao et al., 2019; *Cercidiphyllum japonicum*: Zhu et al., 2020), all have suggested multiple population declines, but sometimes also recoveries, over late Tertiary/Quaternary timescales (Zhu et al., 2020). However, little is known about the relationship between demographic history, genetic variation, long-term survival, and population viability in these Tertiary relict species.

In general, the evolution of deleterious mutation load is a key genetic factor affecting species survival (Butcher 1995; Grossen et al., 2020; Poon & Otto, 2000; Whitlock & Bürger, 2009). Species living in small and isolated populations are expected to experience high levels of inbreeding that will increase the frequency and expression of deleterious recessive mutations (genetic load) (Charlesworth & Willis, 2009). However, levels of genetic load may also be minimized by sustained purging of strongly deleterious mutations resulting from long-term population bottlenecks (Ochoa & Gibbs, 2021). Such a pattern has been observed in several critically endangered species (Benazzo et al., 2017; Grossen et al., 2020; Robinson et al., 2016, 2018; Xue et al., 2015; Yates et al., 2019). Counterintuitively, some Tertiary relict trees do not show signs of such purging but rather indications of mutation load and inbreeding depression (Li et al., 2011; Norén et al., 2016), as evidenced, for example, by poor seedling recruitment in the wild (e.g., *Metasequoia glyptostroboides*: Tang et al., 2011; *Ginkgo biloba*: Tang et al., 2012; Lin et al., 2022).

A large number of Tertiary relict trees are restricted to local habitats with moderate disturbance regimes (Tang et al., 2018). An important factor in the persistence of these small, naturally fragmented populations is almost certainly their capacity for multiple modes of reproduction, including vegetative reproduction that can permit extreme longevity (perhaps over millennia) of locally adapted genets (Bezemer et al., 2019). However, it remains unclear to what extent typical life-history traits of Tertiary relict trees (e.g., individual genet longevity via the ability to repeatedly resprout) and/or population demographic histories have influenced their genome-wide accumulation vs. purging of deleterious mutations (e.g., during periods of population size decline), and thus their local risk of extinction vs. survival and demographic recovery.

Dipteronia Oliv. (Sapindaceae, Hippocastanoideae) contains just two extant deciduous tree species, *D. sinensis* Oliv. and *D. dyeriana* Henry, both endemic to China. This Tertiary relict genus is the closest relative of *Acer* L. (Acevedo-Rodríguez et al., 2010), and is documented in the fossil records of both North America and East Asia from the early Paleocene to the early Oligocene (Ding et al., 2018; Manchester et al., 2009; McClain & Manchester, 2001). The two extant species are similar in several morphological and reproductive characteristics. For instance, individual trees are andromonoecious (having both male and bisexual flowers in paniculate inflorescences), and have winged nutlets dispersed by gravity, wind, and/or water (Xu et al., 2008). Both species reproduce by insect-mediated outcrossing and tend to be more dependent on vegetative reproduction for their recruitment and long-term persistence in disturbed habitats or harsh environments (Qiu et al., 2007). In addition, both trees start bearing fruits at the age of 5–10 years but only reach their full reproductive potential after about 20 years (Ouyang et al., 2006).

Despite these similarities, these two species differ markedly in other key aspects of their biology. *Dipteronia sinensis* and *D. dyeriana* differ in chromosome number, that is, $2n = 2x = 20$ vs. 18 (Löve, 1979; Oginuma et al., 1994; Wolfe & Tanai, 1987), height (*c.* 10–15 vs. 5–10 m), and fruit (nutlet) size (*c.* 2 vs. 5 cm; Appendix S1; Figures S1 and S2) (Ding et al., 2018; Xu et al., 2008). With respect to habitat, *D. sinensis* is widespread in the warm-temperate deciduous forests of central and southwestern China, where it typically grows in sheltered interior forest locations at a wide range of elevations between 1000 and 2400 m a.s.l.; by contrast, *D. dyeriana* is a rare endemic from southeastern Yunnan Province (southwestern China) (Chen, Glémin, & Lascoux, 2017; Xu et al., 2008), where it occurs between 1900 and 2300 m a.s.l. in only a small area (*c.* 6 km^2), within and outside the Wenshan National Reserve (WNR), on the margins of broadleaved evergreen forests (Qiu et al., 2007). Elsewhere in Yunnan and adjacent Guizhou Province, populations of *D. dyeriana* went extinct about 30 years ago (Su et al., 2006; Zhang, 2000). Another natural population located in Malipo County, which is *c.* 100 km away from WNR, was rediscovered in 2017. Surrounded by farmland with less than 10 adult individuals, this small, isolated population is presently at the greatest risk of extinction (Y.X. Qiu, pers. obs.). Where unprotected, the typical forest-margin habitats of *D. dyeriana* have declined considerably in recent decades due to deforestation and clearing for agriculture (Su et al., 2006). Accordingly, differential fluctuations in population size of the two species are likely due to extrinsic factors related to the locations of their ranges, rather than differences in life histories. Overall, *Dipteronia* represents an ideal model system to study the genomic consequences of long-term demographic isolation in an endangered, narrow endemic (*D. dyeriana*) compared to its wider-ranging sister species (*D. sinensis*).

In this study, we first sequenced and *de novo* assembled the chromosome-scale genomes of each *Dipteronia* species. For population genomics analysis, we further sampled 16 populations of *D. sinensis* and seven populations of *D. dyeriana* (Tables S1 and S2). Based on these genomic data, we reconstructed the demographic history of fragmentation and population size reduction of the two species. For intraspecific lineages that appeared to have undergone different demographic histories, we then analyzed patterns of genome-wide diversity and deleterious variation, levels of inbreeding, and the efficiency of purifying selection to gain insights into inbreeding depression and potential risks of population collapse. We addressed the following specific questions: (1) How has past climate change, including Quaternary glacial–interglacial cycles, influenced the demographic histories of each *Dipteronia* species? (2) How has the combination of population bottlenecks and inbreeding shaped the genomic landscapes of each species and their intraspecific lineages? (3) To what extent have deleterious mutations accumulated or been purged in these Tertiary relict species/lineages, and thereby influencing their local extinction risk vs. survival and demographic recovery?

RESULTS

Genome evolution and signatures of positive selection

The *de novo* genomes of *Dipteronia sinensis* and *D. dyeriana* were sequenced to approximately 365× and approximately 250× depth of coverage, respectively (Table S3). The assembled genome sequences were 711.47 Mb (scaffold N50 of 71.32 Mb; contig N50 of 3.82 Mb) in *Dipteronia sinensis* (Figure 1a) and 914.77 Mb (scaffold N50 of 110.51 Mb; contig N50 of 11.27 Mb) in *D. dyeriana* (Figure 1b) with high contiguity, coverage, and accuracy (Appendix S2; Figures S3 and S4; Tables S4–S7). About 61.48% of the *D. sinensis* genome and 65.71% of the *D. dyeriana* genome were composed of repetitive elements (Tables S8 and S9). A total of 33 233 and 32 822 protein-coding genes were predicted in the genomes of *D. sinensis* and *D. dyeriana*, respectively (Table S10).

Phylogenetic analysis of 196 single-copy orthologs shared by all 16 Rosid species included in this study confirmed the sister relationship between *Dipteronia* and *Acer* (Figure 1d) and dated the crown age (the age of the most recent common ancestor) of *Dipteronia* to *c.* 53.84 (95% HPD: 58.15–50.32) million years ago (Mya) (Figure S5). When compared to the genomes of the two *Acer* species for which genome sequences are available, *Dipteronia dyeriana* and *D. sinensis* contained fewer collinear gene pairs (*Dipteronia*: 51.96% vs. *Acer*: 59.37%) and had larger peak values of synonymous substitution rate, K_s (*Dipteronia*: 0.081 vs. *Acer*: 0.045; Table S11, Figure S6a). In the within-species K_s plot (Figure S6b), these four Hippocastanoideae species showed peaks similar to *Vitis vinifera*, indicating that no whole-genome duplications (WGDs) other than the typical core eudicot 'γ event' (Li & Barker, 2020) have occurred in these species.

Using branch-site likelihood-ratio tests (LRTs), we identified 386 putative positively selected genes (PSGs) in *D. sinensis* and 256 PSGs in *D. dyeriana* (Table S12). Among the PSGs in *D. sinensis*, 77 were enriched for the term 'response to stimulus' (GO: 0050896; $P = 3.88e{-}6$, FDR = 0.012; Table S12), including several genes related to cold stress response and plant immunity (e.g., *HPT1*, *DIR7*; Table S13). By contrast, PSGs in *D. dyeriana* mainly encoded genes involved in 'biological regulation' (GO:0065007, 48 genes; $P = 4.08e{-}3$, FDR = 1; Table S12), and more than 20 of them were related to leaf and chloroplast development in response to light (e.g., *CP33*, *ISCA*; Table S13).

Population structure, genomic diversity, and inbreeding

Whole-genome resequencing of 54 *D. sinensis* and 40 *D. dyeriana* individuals generated mean coverages of 24× and 19×, respectively (Table S2). After genotyping and stringent quality filtering, a final set of 5 231 915 SNPs for *D. sinensis* and 2 157 204 SNPs for *D. dyeriana* were selected for genetic diversity estimates (Table S14). Not unexpectedly, the narrow endemic *D. dyeriana* (Figure 2a) had lower genome-wide nucleotide diversity than the wider-ranged *D. sinensis* ($\theta_\pi = 5.15 \times 10^{-4}$, $\theta_W = 4.74 \times 10^{-4}$ vs. $\theta_\pi = 2.98 \times 10^{-3}$, $\theta_W = 1.42 \times 10^{-3}$). Only unlinked and intergenic SNPs (*D. sinensis*: 161503; *D. dyeriana*: 211320) were used for the population structure and phylogenetic analysis. ADMIXTURE identified $K = 2$ and $K = 3$ as the optimal number of genetic clusters for *D. sinensis* and *D. dyeriana*, respectively (Figure S7), consistent with the respective PCAs and ML trees (Figures S8–S10). The two clusters (hereafter 'lineages') of *D. sinensis* were distributed north and south of the Yangtze River, respectively (Figure 2b; North: Qinling-Dabashan Mts. and northeastern Hengduan Mts.; South: eastern Yungui Plateau). Similarly, the three lineages of *D. dyeriana* were also geographically clustered, with two (WS and LZ) in and around the WNR, and one (MLP) in Malipo County (Figure 2b,c). For each species, additional subclusters were detected at higher K values (Figure 2d,e). Notably, despite their smaller geographic range, the three lineages of *D. dyeriana* were more strongly differentiated (pairwise $F_{ST} = 0.22$–0.41) than the North and South lineages of *D. sinensis* ($F_{ST} = 0.04$; Table S15).

Within *D. dyeriana*, all five adult trees examined in MLP were close relatives of less than third-degree, and three of them (MLP01, MLP03, and MLP05) were identified as putatively duplicate samples that may represent clones (ramets) of vegetative propagation (Tables S16 and S17). Closely related individual pairs were also found in the LZ lineage of *D. dyeriana* and the two lineages of *D. sinensis*

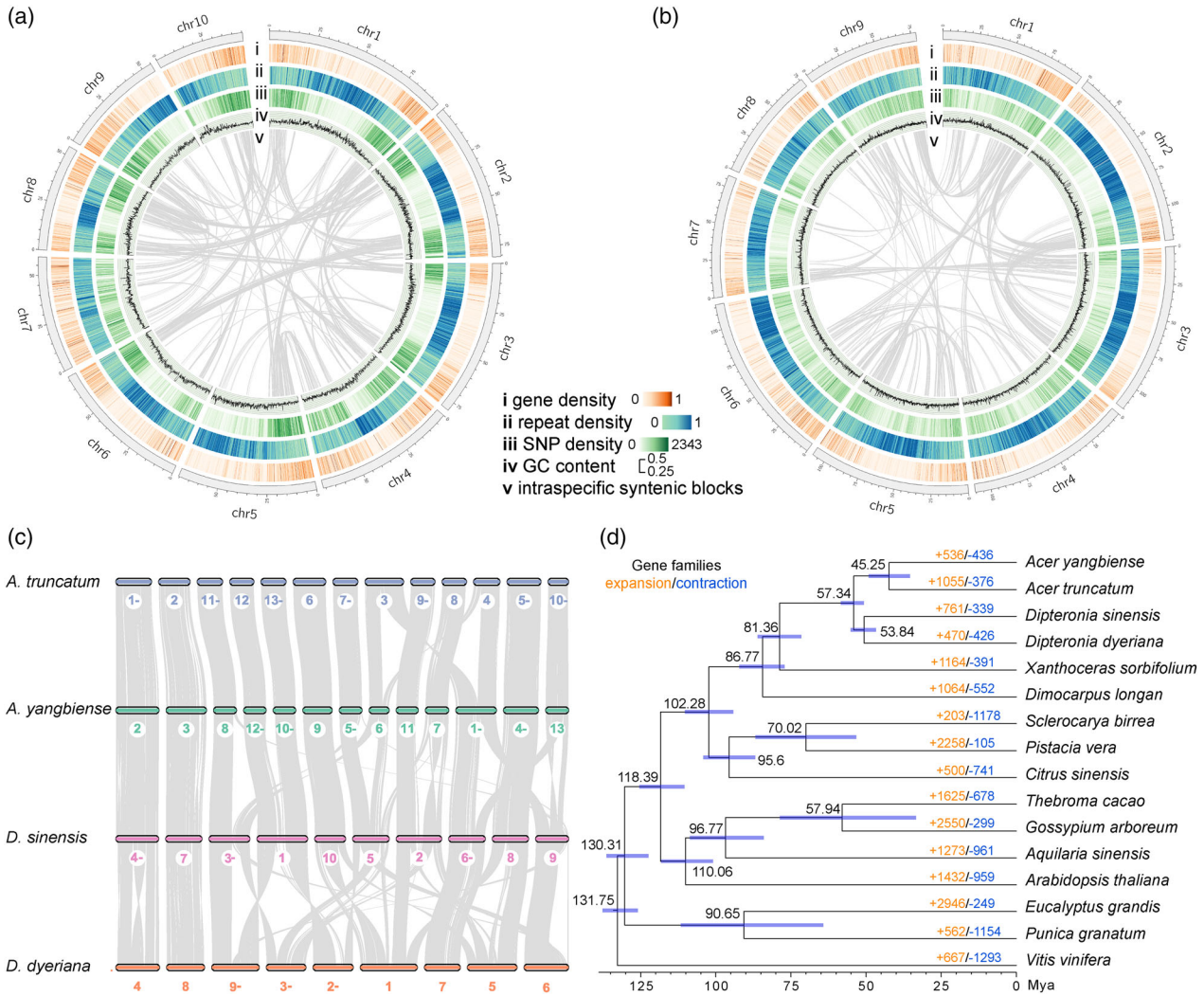

Figure 1. Genome evolution of the two *Dipteronia* species. Circos plots of chromosome-level assembly of the genomes of (a) *D. sinensis* and (b) *D. dyeriana*. The circles, from outermost to innermost, show (i) gene density, (ii) repeat element density, (iii) single nucleotide polymorphism (SNP) density for population genomic analysis, (iv) GC content, and (v) syntenic regions within each genome. The four metrics were calculated in 1 Mb sliding windows.
(c) Synteny pattern among *Acer truncatum*, *A. yangbiense*, *Dipteronia sinensis*, and *D. dyeriana*.
(d) MCMCTREE-derived chronogram of 16 Rosid species, including *D. sinensis* and *D. dyeriana*, based on 196 single-copy orthologous genes. The number at each node represents the median divergence time (in million years ago, Mya), with blue bars representing the 95% highest posterior density (HPD) interval. Numbers in color report the number of gene families that expanded (orange) and contracted (blue) on each branch.

(Tables S16 and S17). In most individuals of each species, we observed a 'sawtooth-like' pattern characterized by long runs of homozygosity (ROH) or low heterozygous regions interspersed with regions of high heterozygosity (Figure S11). Individuals of *D. sinensis* showed, on average, higher genome-scale heterozygosity and lower levels of inbreeding ($H = 0.49$ per kb; $F_{ROH} = 0.30$) than those of *D. dyeriana* ($H = 0.27$ per kb; $F_{ROH} = 0.41$; $P < 0.01$; Figure 3a,b; Figure S12). Within *D. sinensis*, individuals from the North lineage showed on average lower levels of inbreeding ($H = 0.52$ per kb; $F_{ROH} = 0.27$) than those from the South lineage ($H = 0.44$ per kb; $F_{ROH} = 0.36$; $P < 0.05$; Figure 3a,b; Figure S12), which had a smaller inferred effective population size (N_e) (Table S14; see below). Correspondingly, the North lineage had higher genetic diversity ($\theta_\pi = 6.99 \times 10^{-4}$, $\theta_W = 1.27 \times 10^{-3}$) than the South lineage ($\theta_\pi = 5.78 \times 10^{-4}$, $\theta_W = 5.97 \times 10^{-4}$) (Figure 3c; Table S14). In *D. dyeriana*, the MLP lineage had the lowest census population size ($N_c = 5$), followed by LZ ($N_c = 104$) and WS ($N_c = 9671$, Table S1; however, when compared to both latter lineages, MLP showed the lowest level of inbreeding ($H = 0.41$ per kb, $F_{ROH} = 0.18$; $P < 0.01$; Figure 3a,b), as well as the second highest θ_π value (3.53×10^{-4}) and the highest θ_W value (2.95×10^{-4}) (Figure 3c; Table S14).

Within-population estimates of F_{IS} values (window-based) in the two lineages of *D. sinensis* (North, South)

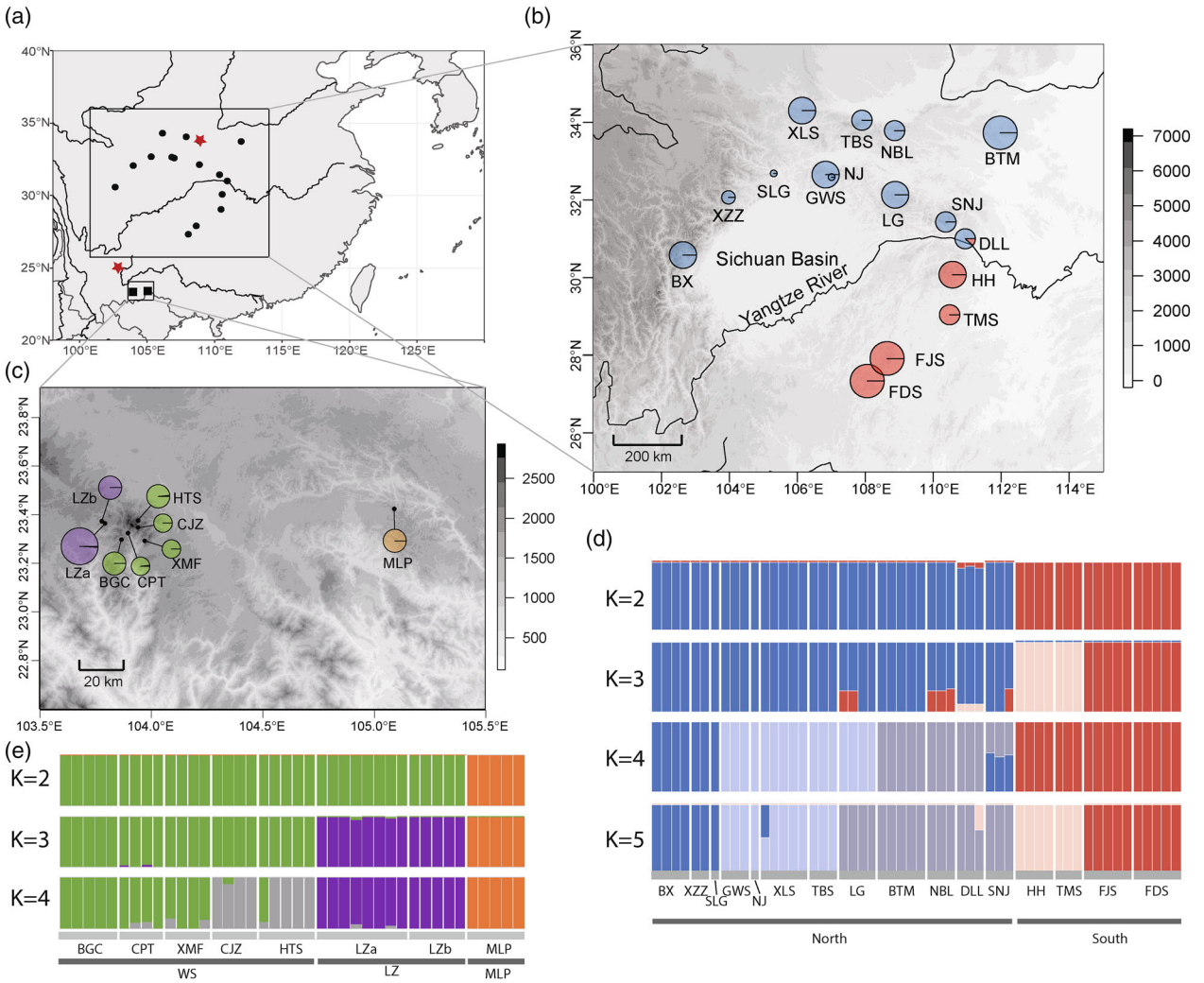

Figure 2. Population genetic structure of the two *Dipteronia* species.
(a) Overall visualization of sampling locations. Dots represent sample locations of *D. sinensis*. The two black rectangles represent the two main distribution locations of *D. dyeriana*. Red stars represent the sample locations of the two reference genome trees.
(b) Geographic distribution of ADMIXTURE clusters for the 15 populations of *D. sinensis* at $K = 2$ (i.e., the optimal solution). Color is in correspondence to (d) and circle size indicates sample size (see Table S2 for detailed sample information).
(c) Geographic distribution of ADMIXTURE clusters for the seven populations of *D. dyeriana* at $K = 3$.
(d) Population genetic structure for *D. sinensis* at $K = 2$ to $K = 5$. North vs. South lineages at $K = 2$ are indicated at the bottom. (e) Population structure for *D. dyeriana* at $K = 2$ to $K = 4$. Names of major lineages at $K = 3$ (i.e., WS, LZ, and MLP) are indicated at the bottom.

and the largest lineage of *D. dyeriana* (WS) were mainly distributed between 0 and 0.5 (Figure 3c; Figure S13), suggesting a degree of inbreeding due to mating between close relatives (Table S17). By contrast, in the smaller-sized LZ and MLP lineages of *D. dyeriana*, F_{IS} values were mostly negative ($F_{IS} < 0$; Figure 3c; Figure S12). Together, the combination of heterozygote excess and high nucleotide diversity detected in the LZ and MLP lineages suggests that the high observed heterozygosity is the legacy of a once much larger, genetically diverse, and outcrossing population, which recently experienced rapid population decline. The increased proportion of vegetative propagation in LZ and MLP may have facilitated the maintenance of high heterozygosity in these small and isolated populations (see below).

Demographic and divergence histories

According to our PSMC analyses of demographic history, most lineages/populations of *D. sinensis* and *D. dyeriana* experienced a prolonged decline in N_e from the late Miocene to the middle Pleistocene (*c.* 5.0–0.8 Mya; Figure 4a,b, Figures S14 and S15); afterward, however, during the Naynayxungla glaciation, *c.* 0.5–0.7 Mya (Zheng et al., 2002), N_e values increased to varying degrees before decreasing again prior to the Last Glacial Maximum (LGM; *c.* 21 thousand years ago, kya). Declines were particularly

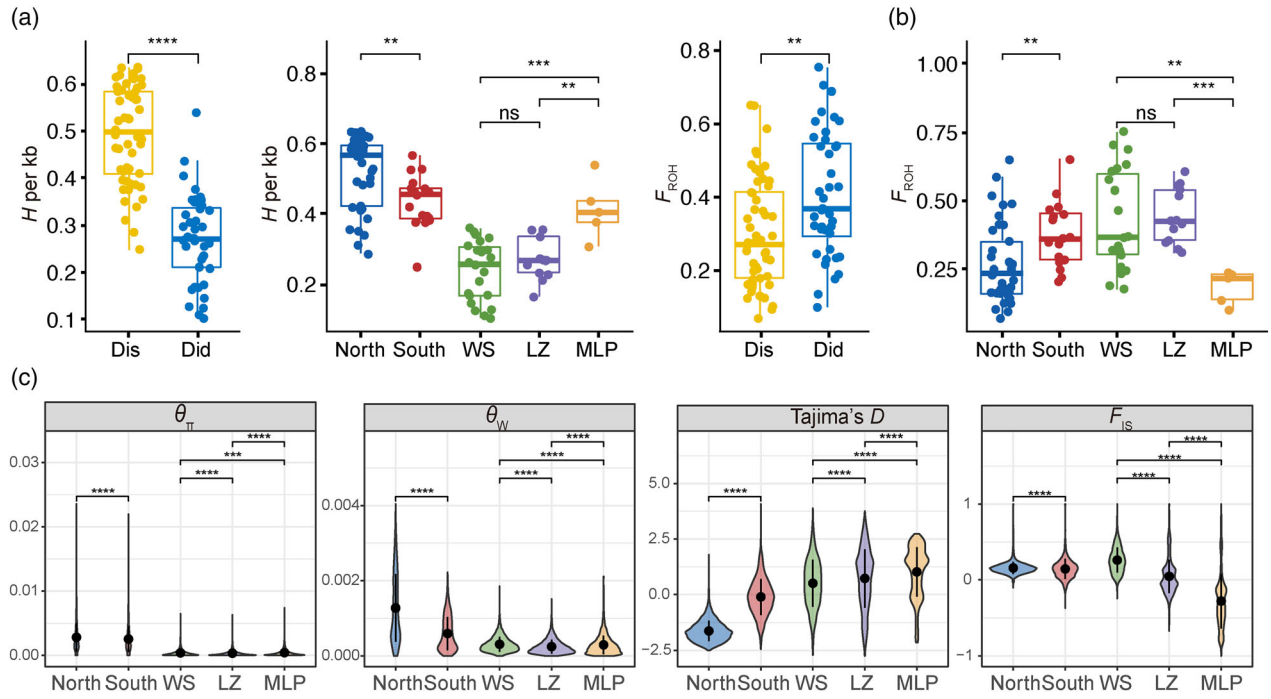

Figure 3. Estimates of inbreeding and genetic diversity. Comparisons of (a) genome heterozygosity (*H*) and (b) estimates of individual genome-based inbreeding (F_{ROH}) between species (Did: *D. dyeriana*; Dis: *D. sinensis*) and lineages (*D. sinensis*, North, South; *D. dyeriana*: WS, LZ, MLP). *P*-values of pairwise comparisons were calculated using Wilcoxon test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, not significant. (c) Distributions of nucleotide diversity parameters (θ_{π}, θ_{W}, Tajima's *D*, and F_{IS}) for each lineage calculated in 100-kb non-overlapping windows across the genome.

pronounced in the two southernmost populations of *D. sinensis* (based on individuals FDS04 and FJS12; see Figure 4a,b, Figures S14 and S15). Considering more recent (i.e., postglacial/Holocene) time scales, SMC++ indicated considerable declines in N_e for most lineages, either slowly (South lineage of *D. sinensis*; WS lineage of *D. dyeriana*) or more rapidly (LZ and MLP lineages of *D. dyeriana*); a notable exception was the North lineage of *D. sinensis*, whose population size increased rapidly after the LGM (Figure S16a,b). STAIRWAY PLOT analyses confirmed the strong recent decline of the MLP lineage (Figure S16c) but indicated relatively stable population sizes for all other lineages during the past 5000 years (Figure S16c).

According to the best-fit FASTSIMCOAL2 model for *D. sinensis* (Figure 4c; Tables S18 and S19), the North and South lineages diverged from their common ancestor in the early Pliocene, *c.* 3.91 Mya. This model also confirmed that the South lineage experienced a contraction after the LGM (Figure 4c; Table S19). For *D. dyeriana*, the best-fit model dated the early divergence of the MLP lineage to *c.* 1.60 Mya and the subsequent split of the WS and LZ lineages to *c.* 0.77 Mya (Figure 4d; Tables S20 and S21). According to this model, both LZ and MLP lineages experienced a massive decrease in N_e over the last 5000 years (Figure 4d; Table S21). For each *Dipteronia* species, estimates of bi-directional gene flow revealed low levels of gene exchange among all pairs of lineages up to the LGM, after which gene flow ceased among lineages (Figure 4c,d).

Accumulation and purging of putatively deleterious mutations

For both *Dipteronia* species, we found weaker efficacy of purifying selection (i.e., higher ratios of π_N/π_S; Kimura, 1983) in smaller-sized populations/lineages compared to larger ones (Table S22). Based on SNPs from protein-coding genes (in total 191 386 vs. 36 693 in *D. sinensis* vs. *D. dyeriana*), we further explored the accumulation vs. purging of putatively deleterious mutations in each species and lineage (Table S23). Among these SNPs, 26 343 (*D. sinensis*) and 4412 (*D. dyeriana*) were predicted to be potentially mildly deleterious mutations by both SIFT and PROVEAN. In addition, SNPEFF identified 2120 and 649 putatively strongly deleterious (putative loss-of-function, LoF) mutations in *D. sinensis* and *D. dyeriana*, respectively. A predicted enrichment in homozygous sites with such deleterious mutations in ROH regions (Szpiech et al., 2013) was observed in only one individual of *D. sinensis* (HH02) and three individuals of *D. dyeriana* (CPT16, MLP03, and MLP04).

For mildly deleterious mutations, statistical analysis of both p_a (proportion of deleterious mutations) and p_g

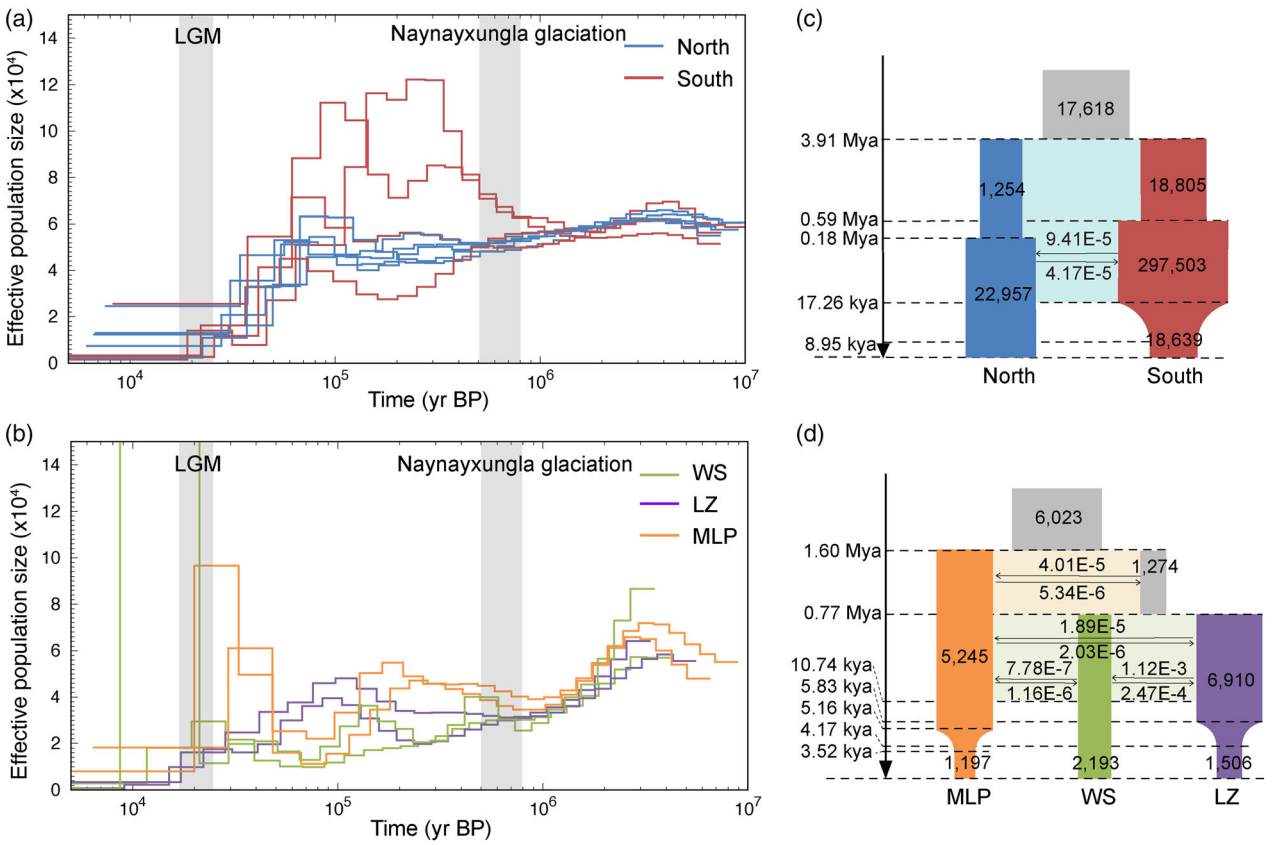

Figure 4. Population size through time plots for (a) *Dipteronia sinensis* (n = 10) and (b) *D. dyeriana* (n = 6) based on pairwise sequentially Markovian coalescent (PSMC) modeling. The time scale on the x-axis is calculated assuming a mutation rate (μ) of 8.17×10^{-9} per site per generation and a generation time (g) of 10 years. The Last Glacial Maximum (LGM; *c.* 21 kya) and the Naynayxungla glaciation (*c.* 0.5–0.7 Mya[17]) are highlighted in gray vertical bars.
(c–d) Demographic scenarios modeled for (c) *D. sinensis* and (d) *D. dyeriana*, using FASTSIMCOAL2, with median times in years, as well as estimates of effective population sizes (N_e) and migration rates (see Tables S19 and S21 for 95% confidence intervals). For each species, estimates of gene flow between linages are given as the migration fraction per generation. kya, thousand years ago; Mya, million years ago.

(proportion of deleterious genotypes, $p_{0/1}$ for heterozygous genotypes and $p_{1/1}$ for homozygous derived genotypes) were significantly higher in *D. dyeriana* than *D. sinensis* ($P < 0.01$; Figure 5a; Table S24). In *D. sinensis*, the South lineage with smaller N_e had significantly higher values of p_a and $p_{1/1}$ ($P < 0.01$) but not $p_{0/1}$ ($P > 0.05$) than the North lineage (Figure 5a; Table S24). In *D. dyeriana*, the smaller-sized MLP lineage showed significantly higher values of p_a and $p_{0/1}$ ($P < 0.05$), but not $p_{1/1}$ compared to LZ and WS lineage ($P > 0.05$, Figure 5a). For the putative LoF mutations, *D. dyeriana* exhibited significantly higher values of p_a and p_g than *D. sinensis* ($P < 0.05$; Figure 5a, Table S24). In *D. dyerina*, the MLP lineage was found to have higher p_a and p_g than the LZ and WS lineages ($P < 0.05$; Figure 5a, Table S24). In *D. sinensis*, the South lineages showed significantly lower p_a and $p_{0/1}$ of LoF mutations than the North lineage ($P < 0.05$), the former had a significantly higher proportion of heterozygous genotypes ($p_{0/1}$) than the latter ($P < 0.05$; Figure 5a,b). After subjecting individuals of each species to linear regression tests, significantly negative correlations were observed between the proportion of mildly deleterious mutations and N_e in *D. sinensis* ($r = -0.43$; $P < 0.01$; Figure 5b). However, no significant correlation was found between proportion of mildly deleterious mutations and N_c or N_e in *D. dyeriana* ($P > 0.05$; Figure 5b). In contrast to mildly deleterious mutations, *D. sinensis* showed no significant correlation between the proportion of LoF mutations and N_e ($P > 0.05$), whereas *D. dyeriana* exhibited a negative correlation between both N_e and N_c and the proportion of LoF mutations ($r = -0.46$ for N_e and $r = -0.73$ for N_c; $P < 0.01$; Figure 5b).

To further assess the relative mutation load between populations and purging of different deleterious mutations in the smaller populations, we used the standardized $R_{X/Y}$ method (Xue et al., 2015) to estimate the relative excess of number of putatively deleterious alleles in one population compared to another (Figure 5c; Table S25). In *D. sinensis*, the smaller-sized South lineage showed an excess of mildly deleterious mutations compared to the larger-sized North lineage ($R_{South/North} = 1.13$), whereas LoF mutations were slightly reduced ($R_{South/North} = 0.98$; Figure 5c;

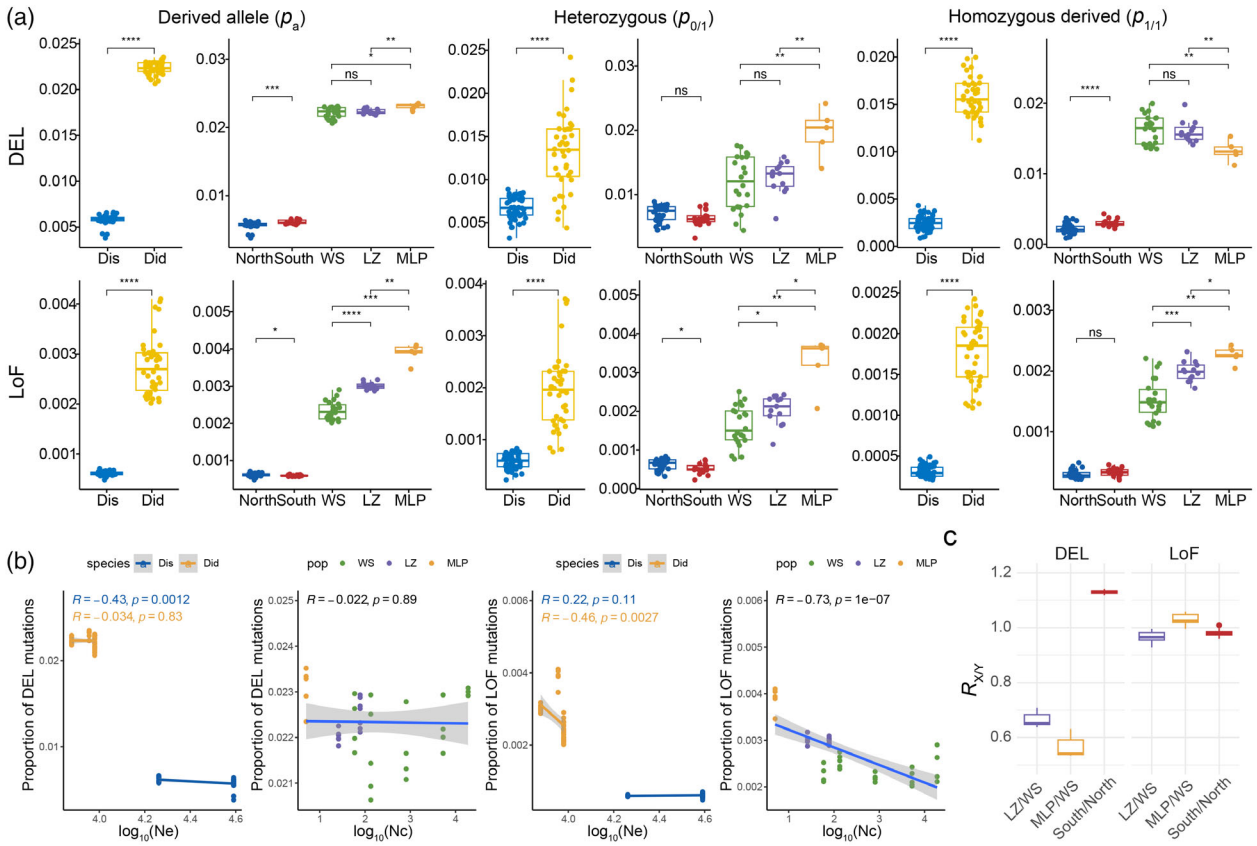

Figure 5. Accumulation and purging of deleterious mutations in the two *Dipteronia* species.
(a) Proportion of derived alleles (p_a), heterozygous genotypes ($p_{0/1}$), and homozygous derived genotypes ($p_{1/1}$) in mildly (DEL) or putatively highly deleterious (LoF) mutations for *D. sinensis* (Dis), *D. dyeriana* (Did), lineages within *D. sinensis* (South, North) and lineages within *D. dyeriana* (WS, LZ, MLP).
(b) Correlation between the proportion of deleterious mutations (DEL, LoF) and effective population size (N_e, see Table S14) across populations/lineages of both species, and between the proportion of mildly deleterious mutations (DEL) and census population size (N_c) across the three populations/lineages of *D. dyeriana*.
(c) Analysis of $R_{X/Y}$ values, contrasting small vs. large populations/lineages in *D. dyeriana* (LZ vs. WS, ML vs. WS) and *D. sinensis* (South vs. North), for DEL and LoF mutations, respectively. $R_{X/Y} < 1$ indicates a relative frequency deficit of mildly deleterious or LoF mutations in small populations compared to large populations. $R_{X/Y}$ distributions are based on jack-knifing across the genome. DEL: single nucleotide polymorphisms (SNPs) with mildly deleterious mutations, as predicted by both SIFT and PROVEAN; LoF: SNPs with strongly deleterious (loss-of-function) mutations, as predicted by SNPEFF. *P*-values of pairwise comparisons were calculated using Wilcoxon test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, not significant.

Table S25). Furthermore, the frequency distributions of LoF mutations, especially in the North lineage, were downward shifted to low frequencies compared to moderate-impact and synonymous mutations (Figure S17). However, in *D. dyeriana*, the smaller-sized LZ and MLP lineages had less mildly deleterious mutations ($R_{MLP/WS} = 0.55$, $R_{LZ/WS} = 0.66$; Figure 5c; Table S25) compared to the larger-sized WS lineage. LZ lineage had less LoF mutations ($R_{LZ/WS} = 0.96$), whereas the MLP lineage had slightly more LoF mutations ($R_{MLP/WS} = 1.02$) compared to the WS lineage. A comparison of scaled additive genetic load revealed more purging of mildly deleterious mutations but less purging of LoF mutations in LZ and MLP than in WS ($P < 0.05$; Figure S18). Moreover, the frequency distribution of LoF mutations was shifted upwards in MLP, with a considerable proportion of those mutations being fixed in this lineage (12.3%; Figure S17).

In the smaller lineage of *D. sinensis* (South), only 6 of 11 genes with fixed LoF mutations, and 20 of 38 genes with fixed mildly deleterious mutations, were annotated successfully (Table S26). In contrast, in the MLP lineage of *D. dyeriana*, 19 of 37 genes with fixed LoF mutations, and 173 of 259 genes with fixed mildly deleterious mutations, were functionally annotated (Table S27). Among these genes, there was some enrichment for particular functions, including stimulus response (e.g., *TPR2*, *AT3G14460*, and *RFNR2*), RNA catabolic process (e.g., *ICME* and *ABCG33*), and seed development (e.g., *PEBP* and *SCRM*) (Table S27).

DISCUSSION

The evolving genomes of two 'living fossil' tree species

We here dated the crown age of *Dipteronia* to the Paleocene/Eocene boundary, *c.* 53.84 Mya (95% HPD: 58.15–

50.32 Mya; Figure 1d), consistent with our previous phylogenomic time estimate (*c.* 55–58 Mya; Feng et al., 2019). This timing broadly coincides with the Paleocene–Eocene Thermal Maximum (PETM), *c.* 55–56 Mya (Schumann et al., 2008; Sexton et al., 2014). In East Asia, the PETM was associated with the transition from semi-arid to more humid conditions due to changes in the monsoon system (Farnsworth et al., 2019). Possibly, therefore, this paleoclimatic event fostered the divergence of the two *Dipteronia* species in southwestern China, for instance, through habitat fragmentation across elevational ranges (Ding et al., 2018).

Consistent with their ancient divergence, *D. sinensis* and *D. dyeriana* markedly differ in genome size (711.47 Mb vs. 914.77 Mb) and chromosome number ($2n = 20$ vs. 18), and exhibit extensive, species-specific genomic rearrangements (Figure 1c). Furthermore, despite a high degree of morphological similarity between the two species (Ding et al., 2018), we identified multiple signatures of species-specific positive selection at functionally important genes related to cold stress response/immunity and leaf/chloroplast development, respectively (Tables S12 and S13). Hence, these potentially adaptive genes may have played a role in the ecological divergence of these two Tertiary relict species.

When compared with the ancient origin of the two *Dipteronia* species, the coalescent times of their extant populations/lineages are relatively young, dating back to either the early Pliocene (*D. sinensis*; Figure 4b) or the early/middle Pleistocene (*D. dyeriana*; Figure 4d; see also Figure S5). Ancient origins of lineages but with recent population coalescent times of extant species are also found in other 'living fossil' tree genera, such as *Ginkgo* (Zhao et al., 2019), *Cercidiphyllum* (Zhu et al., 2020), and *Cycas* (Liu et al., 2021).

Mutation load shaped by historical demography

Our historical demographic models indicate that *D. sinensis* repeatedly recovered from bottlenecks over late Tertiary/Quaternary periods of adverse climate conditions (Figure 4), as also reported for other 'living fossil' trees (e.g., *Ginkgo biloba*: Zhao et al., 2019; *Cercidiphyllum japonicum*: Zhu et al., 2020). By contrast, the effective population size (N_e) of *D. dyeriana* continuously decreased since the late Miocene (Figure 4). Notably, except for the North lineage of *D. sinensis*, all other lineages of both species experienced severe bottlenecks after the LGM (Figure S16), followed by a complete disruption of inter-lineage gene flow (Figure 4c,d). Possibly because of these demographic events, we observed a strong phylogeographic structure in both species (Figure 2; Table S15), with the narrow endemic *D. dyeriana* exhibiting a greater degree of inter-lineage differentiation (WS, LZ, MLP: pairwise $F_{ST} = 0.22$–0.41) compared to the more the widely distributed *D. sinensis* (North, South: $F_{ST} = 0.04$; Table S15). For *D. dyeriana*, our coalescent-based demographic (FASTSIMCOAL2) analyses (Figure 4d) revealed that long-term isolation and/or habitat fragmentation caused by early/middle Pleistocene climate change contributed to prolonged genetic isolation among lineages, regardless of their geographic proximity (especially between LZ and WS).

Moreover, many individuals of both *Dipteronia* species displayed a 'sawtooth-like' pattern of long runs of homozygosity (ROH) interspersed with regions of high heterozygosity (Figure S11). Such isolated and randomly distributed peaks of nucleotide diversity throughout the genome are generally interpreted as signatures of recent inbreeding, as typical for heavily bottlenecked and genetically depleted species (Robinson et al., 2019). Hence, it is feasible that inbreeding and genetic drift, especially in small populations, further promoted lineage divergence in both species (Hartl & Clark, 2007) but also reduced their capacity for demographic recovery (Johnson et al., 2010).

A crucial question arising from this study is how such genetically depleted Tertiary relict tree species could survive, thrive, and successfully compete with taxa occupying a similar ecological niche. Recent empirical studies and population genetic simulations demonstrate the central role of recessive, strongly deleterious mutations in determining extinction risk due to inbreeding depression in small and isolated populations (Grossen et al., 2020; Hedrick & Garcia-Dorado, 2016; Kyriazis et al., 2021; Robinson et al., 2016; Xue et al., 2015). In this study, we generally found a negative and significant ($P < 0.05$) correlation between the proportion of mildly deleterious mutations and population size (N_e) within *D. sinensis* (Figure 5b). This result is not unexpected because purifying selection is more effective in removing mildly deleterious mutations in larger than smaller populations (Glémin, 2003; Welch et al., 2008).

Notably, in *D. sinensis*, the proportion of loss-of-function (LoF) mutations showed no significant correlation with N_e (Figure 5b). The smaller South lineage even had relatively fewer LoF mutations than the larger North lineage (Figure 5a). This difference is likely related to the contrasting demographic histories of the two lineages. According to our FASTSIMCOAL2 and SMC++ analyses, the North lineage experienced recent population growth after the LGM (Figure 4b; Figure S16), likely because of postglacial northward expansion of *D. sinensis* from south of the Yangtze to the Daba-Qinling Mountains (Harrison et al., 2001). In contrast, the South lineage underwent a population bottleneck *in situ* during postglacial/Holocene periods (Figure S16a) and showed higher levels of individual inbreeding (Figure 3a,b). Hence, in the smaller South lineage of *D. sinensis*, LoF variants were likely exposed more frequently in a homozygous state as observed in Figure 5a, and were therefore purged by inbreeding and purifying selection (see also Glémin, 2003; Xue et al., 2015). This inference is further supported by the $R_{X/Y}$ ratios (Figure 5c) and stands in

contrast to observations made in the smallest population MLP of *D. dyeriana* (discussed below), where LoF variants apparently accumulated rather than being purged.

Living on the verge of extinction

Concern about the survival of 'living fossil' trees, and critically endangered species more generally, is focused largely on the threat from increased anthropogenic pressure. If population sizes have decreased to a threshold size, recovery will be constrained and populations (and species) may enter an 'extinction vortex' (Hutchings, 2015; Yang et al., 2018). This, in particular, could be the case for the LZ and MLP lineages of *D. dyeriana*, which are presently located outside of the Wenshan National Reserve (WNR) and thus exposed to human disturbance (Zhang, 2000; see also Appendix S1. Based on our FASTSIMCOAL2 and SMC++ analyses, the LZ and MLP lineages experienced a more massive and serious decrease in N_e than WS over the last 5000 years (Figure 4d; Figure S16, Table S21). Likely as a consequence, MLP currently has the smallest range and population size among all three lineages of *D. dyeriana*, but still harbors higher levels of nucleotide diversity and genomic heterozygosity than LZ and WS (Table S14; Figure 3a). Previous studies have suggested that vegetative reproduction and the long-lived nature of the clones can mitigate the impact of genetic drift, preventing the loss of genetic variation in smaller populations (Gargiulo et al., 2018; Young et al., 1996). This also seems to explain why small-sized lineages of Tertiary relict trees (LZ, MLP) can still retain considerable heterozygosity (MLP), and thus despite historical population decline (see also Robinson et al., 2016).

The LZ and MLP lineages were found to harbor more recessive DEL mutations as compared to WS lineage (Figure 5c; Figure S18), but have relatively smaller $R_{X/Y}$ values. This argues for a mechanism whereby these highly recessive DEL mutations can be removed by strong genetic drift in these populations experiencing recent severe bottlenecks (Glémin, 2003). Nonetheless, it is noteworthy that the total number of DEL in LZ and MLP was still equal to or greater than that in the WS lineage (Figure 5a). By contrast, despite exhibiting a degree of purging of strongly deleterious mutations in LZ and MLP as manifested by $R_{X/Y}$ values (Figure 5c; Figure S18), the two smaller-sized lineages of *D. dyeriana* harbored substantially higher proportions of LoF mutations and larger proportions of fixed LoF mutations as compared to the larger-sized WS lineage located within WNR (Figure 5a). This suggests that partial purging of strongly deleterious mutations was not sufficient to erase a substantial fraction of this type of mutation load due to reduced efficacy of purifying selection in the small–isolated populations (Glémin, 2003; Khan et al., 2021). In fact, this pattern mirrors those found in other threatened species that have experienced recent severe population bottlenecks (*Canis lupus*: Robinson et al., 2019; *Cyprinodon diabolis*: Tian et al., 2022); and it clearly contradicts the notion that strongly deleterious mutations are more likely to be purged from small, bottlenecked populations (Hedrick & Garcia-Dorado, 2016; Robinson et al., 2018; Yang et al., 2018). Thus, for *D. dyeriana*, our mutation load results and negative correlations between population size (N_e and N_c) and LoF mutations (Figures 5a,b) generally support the theory prediction that threatened species living in small populations will have a high mutation load (van Oosterhout, 2020).

Nonetheless, any further decline (e.g., due to human disturbance and climate change) would cause strongly deleterious recessive alleles (carried as heterozygotes in a historically large population) to quickly become homozygous through inbreeding, thereby exacerbating the severity of inbreeding depression (Charlesworth & Willis, 2009; Kyriazis et al., 2021). According to our field observations, seedling establishment is surprisingly low in the LZ and MLP populations of *D. dyeriana* (Table S1) (Qiu et al., 2007). Moreover, the higher proportion of clonal individuals genetically inferred for LZ and MLP than WS (Table S17) indicates an increased proportion of vegetative propagation in small and isolated populations of this species, as reported for other 'living fossil' trees (e.g., Lin et al., 2022; Tang et al., 2011, 2012). While providing a short-term evolutionary solution to achieving reproductive assurance in unfavorable environmental conditions (Lindborg & Eriksson, 2004), a shift towards increased vegetative propagation, combined with long generation times in these Tertiary relict trees, may hinder the purging of strongly deleterious mutations; this in turn can result in a rapid accumulation of deleterious alleles throughout the genome, ultimately trapping small isolated populations in an extinction vortex (Lynch et al., 1995; McKey et al., 2010; Ramu et al., 2017). In line with this assumption, our results show that increased accumulation of LoF mutations in MLP and LZ, possibly due to increased genetic drift and reduced selection efficacy, has resulted in the fixation of LoF mutations in functionally important genes related to stimulus response, RNA catabolic process and seed development (Figure S17, Table S27).

Although our estimates of mutation load do not directly measure fitness, there is empirical evidence that LoF mutations are, on average, deleterious (Eyre-Walker & Keightleyet al., 2007), which may have severely deleterious consequences for the possible future survival of *D. dyeriana*, as recently demonstrated for other organisms (e.g., inbred gray wolf: Gómez-Sánchez et al., 2018; Robinson et al., 2019; Devils Hole pupfish: Tian et al., 2022). Therefore, given all this evidence, the currently unprotected MLP and LZ lineages of *D. dyeriana* should have top conservation priority, as both represent separate evolutionary significant units (ESUs) (Moritz, 1994), and appear to have a

greater mutation load than the protected WS lineage. Moreover, rather than just maintaining extant population sizes of LZ and MLP, future efforts should focus on their genetic rescue by reforesting them with individuals screened for low genetic load, for example, from WS.

CONCLUSIONS

The two extant species of *Dipteronia* evolved independently for more than 50 million years with little morphological change but experienced different demographic histories since the late Miocene. Our genome-scale analyses of each species are the first to provide insights into the genomic landscape of deleterious variation in Tertiary relict trees. Our results suggest that the widespread *D. sinensis* repeatedly recovered from demographic bottlenecks over late Tertiary/Quaternary periods of adverse climate conditions, while the population size of the narrow endemic *D. dyeriana* has steadily decreased since the late Miocene, and especially after the LGM. The more efficient purging of deleterious mutations in *D. sinensis* may have facilitated its survival and repeated demographic recovery. By contrast, in *D. dyeriana*, increased genetic drift and reduced selection efficacy, due to recent drastic population declines and a likely preponderance of vegetative propagation, resulted in the fixation of strongly deleterious mutations (especially in functionally important genes of the two smaller lineages LZ and MLP), reduced fitness and continuous population decline. These effects likely have detrimental consequences for the species' future viability and adaptive potential.

Given the large number of deleterious variations in *D. dyeriana*, as likewise observed in other small and isolated populations of endangered species (Tian et al., 2022), management strategies could prioritize minimizing deleterious variation in order to potentially maximize mean absolute fitness (van Oosterhout, 2020). With this approach, genomic data should be used to identify individuals with the fewest number of strongly deleterious variants from the large source population for genetic rescue (Kyriazis *et al.*, 2021). Moreover, in future conservation efforts, careful predictions of putative phenotypic effects of highly deleterious variants could be instructive in understanding future inbreeding depression, and these variants could potentially be used as 'early warning signs' for population decline (Khan et al., 2021). Taken together, our results thus demonstrate that contrasting demographic histories have different effects on the extent of accumulation and purging of putatively deleterious mutations that likely determine the long-term survival and extinction risk of Tertiary relict trees. More broadly, these insights have important implications for the genetic management of deleterious variants in small and isolated populations of Tertiary relict trees, and flowering plants in general, in the increasingly fragmented landscape of the modern world.

EXPERIMENTAL PROCEDURES

Genome sequencing, assembly, annotation, and comparative genomic analyses

We collected fresh young leaves from one adult tree each of *Dipteronia sinensis* and *D. dyeriana* in the Niubeiliang National Natural Reserve (Shaanxi, China) and Kunming Botanical Garden (Yunnan, China; transplanted from Wenshan National Reserve, WNR), respectively. We sequenced the genome of each individual using a combination of PacBio single-molecule long reads and Illumina short reads and further improved the assembly by adding 10× Genomics linked-reads. Then, Hi-C libraries were sequenced to assist the chromosome-level assembly. Gene prediction was conducted through a combination of homology-based inference, *ab initio* prediction, and transcripts from RNA sequencing. A combined strategy based on homology alignment and *de novo* were also applied in our repetitive element annotation pipeline (see details in Appendix S2).

The two genomes of *Dipteronia* (*D. sinensis* and *D. dyeriana*) and those of two other members of Hippocastanoideae, *Acer yangbiense* (Yang et al., 2019) and *A. truncatum* (Ma et al., 2020), were included in the genome syntenic analysis. A mutation rate for this subfamily was calculated from synonymous, four-fold degenerate sites (FFS) in single-copy orthologous sequences shared by the two *Dipteronia* species and the two *Acer* species. *Vitis vinifera* (Vitaceae), which lacks a recent whole-genome duplication (WGD) event (Jaillon et al., 2007), was included to infer whether the two *Dipteronia* genomes have undergone recent WDG events. For our time-calibrated phylogenetic reconstructions, including analyses of gene family evolution, we included the above five species plus 11 others from the Rosid clade of angiosperms (Folk et al., 2018). Positively selected genes (PSGs) in each *Dipteronia* species were identified by branch-site likelihood-ratio tests (LRTs) using the program CODEML from the PAML package v4.9 (Yang, 2007). Gene ontology (GO) enrichment analysis was conducted on the corresponding orthologous genes in *Arabidopsis thaliana*. See detailed methods of this section in Appendix S2.

Field studies, population sampling, and whole-genome resequencing

Individual leaf samples were collected from seven populations (n = 40) of *D. dyeriana* and 16 populations (n = 54) of *D. sinensis* (Tables S1 and S2). In populations with more than 10 adult individuals, samples were taken more than 10 meters apart; otherwise, all adult individuals were sampled (see Appendix S1 for details). For *D. dyeriana*, our field studies (Appendix S1) showed that the seven populations of *D. dyeriana* covered 545 716 m^2, with a total census size (N_c) of 9780 adult trees, and with individual populations ranging from 1013 m^2 (MLP, N_c = 5) to 236 113 m^2 (CJZ, N_c = 6737; Table S1).

For whole-genome resequencing of all 94 individuals, TRIMMOMATIC v0.39 (Bolger et al., 2014) was used to remove adapter sequences, potential contamination, and low-quality bases. All of the trimmed reads of each individual were then mapped to their corresponding reference genomes using BOWTIE v2.2.5 (Langmead & Salzberg, 2012). Duplicate alignments were marked using PICARD

tools v1.1 (http://broadinstitute.github.io/picard/). GATK HAPLOTYPECALLER v3.5 was used to call variants. Identified single nucleotide polymorphisms (SNPs) were filtered by "QD < 5.0 || FS > 10.0 || MQ < 40.0 || SOR > 2 || MQRankSum < −3.0 || MQRankSum > 3.0 || ReadPosRankSum < −2.0 || ReadPosRankSum > 2.0 || Inb > −0.4 || EH < 5" with additional filtering steps including removing: (1) SNPs with more than two alleles; (2) genotypes with average quality scores (mean GQ) < 20, or extremely low (< one-third average depth) or extremely high (> 10 average depth) coverage; and (3) SNPs with more than half missing genotypes.

Population structure, genome-wide genetic diversity, and inbreeding

For each *Dipteronia* species, population structure was assessed by Bayesian clustering in ADMIXTURE v1.23 (Alexander et al., 2009) based on unlinked and putatively neutral (i.e., intergenic) SNPs. The number of genetic clusters (K) was set to vary from 2 to 10, with the optimal K selected by the lowest cross-validation error. To further examine population genetic relationships, we subjected the above SNP dataset to both a principal component analysis (PCA) in GCTA v1.24 (Yang et al., 2011) and a maximum likelihood (ML) tree analysis in RAXML v8 (Stamatakis 2014). For each species and intraspecific lineage (as identified by ADMIXTURE), we further calculated polymorphism statistics in PIXY v1.0.0 (Korunes & Samuk, 2021) and VCFTOOLS v0.1.16 (Danecek et al., 2011) along the genome in 100-kb non-overlapping windows, including nucleotide diversity (θ_π), Watterson's estimator (θ_W), and Tajima's D. Contemporary effective population size (N_e) for each species and lineage was calculated using the formua: $N_e = \frac{\theta_W}{4\times\mu}$ (Wright, 1931). In addition, we used KING v2.2.5 (Manichaikul et al., 2010) to assess the genetic relatedness (kinship coefficient) between individuals of each lineage. Finally, we used VCFTOOLS v0.1.16 (Danecek et al., 2011) and PLINK v1.9 (Purcell et al., 2007) to calculate individual, genome-based coefficients of homozygosity or 'autozygosity' (in terms of F_{ROH}) and heterozygosity (H), as well as 'classical' individual- and population-based estimates of inbreeding (F_{IS}) (see Appendix S3 for details).

Demographic inference and simulations

Because of their relative advantages and limitations, we used four complementary methods to infer the demographic history of each *Dipteronia* species, based on the re-sequencing data of nine individuals of *D. sinensis* and six of *D. dyeriana*, representing each regional lineage per species. First, we used the pairwise sequentially Markovian coalescent (PSMC) method (Li & Durbin, 2011) to infer historical changes in N_e based on representative individuals of each lineage per species. However, the PSMC approach should be treated with caution when dealing with evolutionary histories more recent than 20–30 thousand years ago (Schiffels & Durbin, 2014), or when coverage is low (Nadachowska-Brzyska et al., 2016). Therefore, we also applied the sequentially Markovian coalescent (SMC) method, implemented in SMC++ v1.13 (Terhorst et al., 2016), as well as the program STAIRWAY PLOT v2.0 (Liu & Fu, 2020) to estimate recent changes in N_e with greater accuracy (Liu & Fu, 2020). Finally, we employed coalescent simulations in FASTSIMCOAL2 v2.6 (Excoffier et al., 2013), using unfolded joint site frequency spectra (u-SFS) generated from unlinked intergenic SNPs. The scenarios and corresponding prior distributions of the FASTSIMCOAL2 parameters (e.g., lineage divergence, lineage divergence times, N_e) were set according to results from ADMIXTURE and the above demographic analyses (e.g., PSMC; see Appendix S4 for details).

Estimates of selection efficacy and mutation load

To evaluate selection efficacy for *D. sinensis* and *D. dyeriana*, we estimated nucleotide diversity (π) at nonsynonymous (zero-fold) and synonymous (four-fold) sites (i.e., π_N and π_S, respectively) for each species and lineage using an in-house PERL script (Chen, Lu, et al., 2017). To estimate their genome-wide levels of mutation load, we classified the SNPs of both *D. sinensis* and *D. dyeriana* into synonymous, nonsynonymous, and putative loss-of-function (LoF) mutations, using SNPEFF v4.3 (Cingolani et al., 2012). The effect of the nonsynonymous SNPs was assessed using SIFT4G (Vaser et al., 2016) (database: UniRef90) and PROVEAN v1.1.5 (Choi et al., 2012) (database: NR). Positions with a SIFT score <0.05 and PROVEAN score < −2.5 were predicted to be deleterious (hereafter 'mildly deleterious mutations'). To mitigate reference bias, we included *Acer yangbiense* and *A. truncatum* (see above) for the purpose of polarizing variable sites into ancestral and derived states. Specifically, the two *Acer* species and *D. sinensis* served as outgroups for *D. dyeriana*, while both *Acer* species and *D. dyeriana* were utilized as outgroups for *D. sinensis*. We filtered SNPs under the additional two criteria for the subsequent analyses: they had no missing genotypes, and they exhibited consistent homozygous states across all three outgroup species. We used EST-SFS v2.03 to infer the probability of the derived versus ancestral allelic state, with a probability cutoff set at 0.95, as described by Keightley and Jackson (2018). The ancestral allele was set as the non-deleterious allele, and the derived allele as the potential deleterious allele.

We calculated deleterious mutation load using the following two indexes. The proportion of deleterious mutations in each category was calculated as: $p_a = \frac{N^{class}_{derived\ alleles}}{N^{total}_{SNPs}\times 2}$, where $N^{class}_{derived\ allele}$ is the number of derived alleles (counting heterozygotes once and homozygotes twice) in each category (mildly deleterious and LoF) of SNPs, and N^{total}_{SNPs} is the number of SNPs used in the annotation of deleterious mutations (see Results). The proportion of heterozygous ($p_{0/1}$) and homozygous ($p_{1/1}$) derived genotypes in each category was calculated as: $p_g = \frac{N^{class}_{genotype}}{N^{total}_{SNPs}}$. Values of p_a and p_g were calculated for each individual and compared among species and lineages through Wilcoxon tests. Besides the above statistics, we also scaled the numbers of derived alleles/genotypes of DEL and LoF SNPs by the corresponding numbers of synonymous SNPs to evaluate the efficacy of purifying selection at the individual level (scaled additive load, González-Martínez et al., 2017). To further test for differences in the purging of deleterious mutations between populations, we calculated the expected number of derived mutations at each variable site in one genome sampled from population X that are not seen in a randomly selected genome from population Y and vice versa; and we calculated a ratio of the two counts: $R_{X/Y}$ (Do et al., 2015). Following Do et al. (2015), we used a random set of putatively neutral SNPs (i.e., intergenic SNPs) for standardization, which makes $R_{X/Y}$ robust against sampling effects and population substructure. For a specific set of putative deleterious sites, $R_{X/Y} > 1$ denotes excess of derived deleterious mutations in population X compared to population Y, while $R_{X/Y} < 1$ means deficiency of derived deleterious mutations (i.e., population X has stronger purging of deleterious mutations than population Y). For each species, we calculated $R_{X/Y}$ for small (X) vs. large (Y) populations, with population size determined by N_e calculated for lineages of both species and N_c of the investigated populations of *D. dyeriana*. Finally, to assess whether deleterious mutations in small populations led to a reduction in fitness, we used homology-based inferences to functionally annotate genes with fixed deleterious mutations in the South lineage of *D. sinensis* and the MLP lineage of *D. dyeriana*.

ACCESSION NUMBERS

All sequencing data used in this study have been deposited in the NCBI SRA database with the Bioproject number PRJNA796066. These reference genomes and gene annotations have been uploaded to the NCBI Genome database under accessions JANJYI000000000 and JANJYJ000000000 and China National GeneBank DataBase (CNGBdb) under the accession CNP0002617. The custom scripts used in this study have been uploaded to the GitHub repository (https://github.com/Fengyaa/Dipteronia_genome/).

ACKNOWLEDGEMENTS

We thank Xinxin Zhu, Shuai Liao, Xinpeng Zhou, Guifeng Cui, Shaanxi Niubeiliang National Natural Reserve Management Bureau, Dalaoling National Nature Reserve Management Bureau, Baoxing County Forestry Bureau, Houhe National Nature Reserve Management Bureau, Wenshan National Nature Reserve Management Bureau, Malipo County Forestry Bureau, and Kunming Institute of Botany for assistance with sampling; and Dr. Tao Ma from Sichuan University for valuable comments and discussions. We also gratefully acknowledge the valuable comments by three anonymous reviewers on a previous version of this manuscript. This work was supported by the National Key R&D Program of China (grant Nos. 2022YFF1301703, 2017YFA0605100), the National Natural Science Foundation of China (grant Nos. 31872652 and 32161143003), and the Science and Technology Basic Resources Investigation Program of China (grant No. 2017FY100100).

CONFLICT OF INTEREST

All authors confirm that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

YQ conceived the project. YQ, YF, and KMO designed the experiments and coordinated research activities. YF, RL, PL, and SZ collected samples. YF, JC, XZ, and SZ analyzed data. KMO and JQ contributed to the interpretation of results. YF, YQ, KMO, and HPC wrote and revised the manuscript. All authors read and approved the manuscript.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Morphological characteristics and habitat of *D. dyeriana*.

Figure S2. Morphological characteristics and habitat of *D. sinensis*, and comparison of the fruits (schizocarps) of the two extant and fossil species of *Dipteronia*.

Figure S3. 17-mer-based analysis to estimate the genome sizes of *D. sinensis* and *D. dyeriana*.

Figure S4. The Hi-C chromatin interaction map of 10 pseudo-chromosomes for *D. sinensis* and nine pseudo-chromosomes for *D. dyeriana*.

Figure S5. MCMCTREE-derived chronogram of 16 Rosid species.

Figure S6. Distribution of synonymous substitution rate (K_s) for collinearity blocks between species pairs and within species for *Dipteronia sinensis*, *D. dyeriana*, *Acer yangbiense*, *A. truncatum*, and *Vitis vinifera*.

Figure S7. Cross-validation error values of different number of clusters (*K*) set in the ADMIXTURE analysis.

Figure S8. Principal component analysis (PCA) based on unlinked intergenic single nucleotide polymorphisms (SNPs) for individuals.

Figure S9. Maximum likelihood (ML) tree for 54 individuals of *D. sinensis* based on unlinked SNPs.

Figure S10. Maximum likelihood (ML) tree for 25 individuals of *D. dyeriana* based on unlinked SNPs.

Figure S11. Distributions of heterozygosity across the genome of 25 individuals of *D.* dyeriana and 54 individuals of *D. sinensis*.

Figure S12. Per-site genome-wide heterozygosity, inbreeding, and F_{ROH}.

Figure S13. Distribution of nucleotide diversity (π), Tajima's *D*, and inbreeding coefficient (F_{IS}) across the genome presented in non-overlapping 1 Mb windows.

Figure S14. Changes in historical population sizes (N_e) inferred by the pairwise sequentially Markovian coalescent (PSMC) method for representative individuals of *D. sinensis*.

Figure S15. Changes in historical population sizes (N_e) inferred by the pairwise sequentially Markovian coalescent (PSMC) method for representative individuals of *D. dyeriana*.

Figure S16. Estimation of effective population size change through time for each *Dipteronia* lineage using SMC++ (a,b) and STAIRWAY PLOT (c).

Figure S17. Site frequency spectrum (SFS) for neutral (synonymous, SYN), mildly deleterious (DEL), and highly deleterious (LoF) mutations for each lineage of *D. dyeriana* (WS, LZ, and MLP) and *D. sinensis* (South, North).

Figure S18. Numbers of derived mutations, heterozygous genotypes, and homozygous genotypes in DEL and LoF SNPs, scaled by the corresponding numbers in synonymous SNPs.

Figure S19. Visualization of observed and simulated 2D-SFS (two-dimensional site frequency spectrum) and marginal 1D-SFS (one-dimensional site frequency spectrum) for the best scenario of *D. sinensis* (a) and *D. dyeriana* (b–d).

Figure S20. FASTSIMCOAL2 scenarios for testing the existence and direction of gene flow between pairs of populations of *D. sinensis* and *D. dyeriana*, respectively.

Figure S21. FASTSIMCOAL2 scenarios used for testing the divergence and demographic history of the two lineages of *D. sinensis*.

Figure S22. FASTSIMCOAL2 scenarios used for testing the divergence and demographic history of the three lineages of *D. dyeriana*.

Figure S23. Schematic diagram of the best-fitting divergence and demographic history model for *D. sinensis* (a) and *D. dyeriana* (b), as inferred by FASTSIMCOAL2.

Figure S24. Distribution of nucleotide diversity calculated from SNPs filtered by different criteria.

Table S1. Population investigation and sampling information for *D. dyeriana*.

Table S3. Sequencing data of *D. sinensis* and *D. dyeriana*.

Table S4. Genome assembly results using methods combining PacBio single-molecule long reads, Illumina short reads, 10× Genomics Linked-Reads, and Hi-C technology.

Table S5. Genome assembly completeness evaluation by Core Eukaryotic Genes Mapping Approach (CEGMA) for the two *Dipteronia* species.

Table S6. Genome assembly completeness evaluation by Benchmarking Universal Single-Copy Orthologs (BUSCO) for the two *Dipteronia* species.

Table S7. Hi-C chromosomal mapping statistics.

Table S8. Repeat elements in the genomes of the two *Dipteronia* species, as predicted by different tools.

Table S9. Classification statistics of repeat elements for *D. sinensis* and *D. dyeriana*.

Table S10. Statistical comparison of structure annotation for genomes of the two *Dipteronia* species and seven closely related species.

Table S11. Collinearity statistics for species pairs of *Acer* and *Dipteronia*.

Table S14. SNP calling and genetic diversity statistics for each species and lineage.

Table S15. Estimation of population divergence (F_{ST} values) between lineages of *D. sinensis* and *D. dyeriana*, respectively.

Table S16. Average kinship coefficient between individual pairs from each sampled population, and the proportion of closely related pairs per *Dipteronia* lineage.

Table S18. Likelihood comparison of FASTSIMCOAL2 models for divergence history and population size change of the two lineages of *D. sinensis*.

Table S19. Parameter estimates with 95% highest posterior density (HPD) intervals for the best FASTSIMCOAL2 model of *D. sinensis* with parameter tags corresponding to Figure S23a.

Table S20. Likelihood comparison of FASTSIMCOAL2 models for divergence history and historical population size change of the three lineages of *D. dyeriana*.

Table S21. Parameter estimates with 95% highest posterior density (HPD) intervals for the best FASTSIMCOAL2 model of *D. dyeriana* with parameter tags corresponding to Figure S23b.

Table S22. Nucleotide diversity at nonsynonymous (π_N) and synonymous (π_S) sites for the two *Dipteronia* species and lineages.

Table S23. Number of total SNPs in the coding regions and potential deleterious SNPs used to evaluate mutation load of each *Dipteronia* species.

Table S24. Mean proportion of derived alleles (p_a), heterozygous genotypes ($p_{0/1}$), and homozygous derived genotypes ($p_{1/1}$) in each class of deleterious sites for the two *Dipteronia* species and their lineages.

Table S25. $R_{X/Y}$ values, contrasting small vs. large populations/lineages in *D. dyeriana* (LZ vs. WS, ML vs. WS) and *D. sinensis* (South vs. North), for each class of mutations.

Appendix S1. Field investigations of the two *Dipteronia* species.

Appendix S2. Genome sequencing, assembly, and annotation.

Appendix S3. Estimate of genomic diversity and inbreeding.

Appendix S4. Lineage divergence and demographic history of the two *Dipteronia* species.

Table S2. Sampling information and sequencing statistics for all individuals of the two *Dipteronia* species.

Table S12. Gene Ontology (GO) term enrichment analysis for positively selected genes (PSGs) identified in the two *Dipteronia* species.

Table S13. Positively selected genes (PSGs) identified in the two *Dipteronia* species and orthology-based annotation retrieved from the TAIR database (https://www.arabidopsis.org/).

Table S17. Kinship of individual pairs sampled in the same location inferred by the KING program.

Table S26. Genes with fixed deleterious mutations in the South lineage of *D. sinensis* and functional annotations of their orthologous genes in *Arabidopsis thaliana*.

Table S27. Genes with fixed deleterious mutations in the MLP lineage of *D. dyeriana* and functional annotations of their orthologous genes in *Arabidopsis thaliana*.

REFERENCES

Acevedo-Rodríguez, P., **van Welzen, P.C.**, **Adema, F.** & **van der Ham, R.W.J.M.** (2010) Sapindaceae. In: Kubitzki, K. (Ed.) *Flowering plants. Eudicots: Sapindales, Cucurbitales, Myrtaceae*. Berlin, Germany: Springer, pp. 357–407.

Alexander, D.H., **Novembre, J.** & **Lange, K.** (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**, 1655–1664.

Benazzo, A., **Trucchi, E.**, **Cahill, J.A.**, **Delser, P.M.**, **Mona, S.**, **Fumagalli, M.** ***et al.*** (2017) Survival and divergence in a small group: the extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proceedings of the National Academy of Sciences, USA*, **114**, E9589–E9597.

Bezemer, N., **Krauss, S.L.**, **Roberts, D.D.** & **Hopper, S.D.** (2019) Conservation of old individual trees and small populations is integral to maintain species' genetic diversity of a historically fragmented woody perennial. *Molecular Ecology*, **28**, 3339–3357.

Bolger, A.M., **Lohse, M.** & **Usadel, B.** (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Butcher, D. (1995) Muller's ratchet, epistasis and mutation effects. *Genetics*, **141**, 431–437.

Charlesworth, D. & **Willis, J.H.** (2009) The genetics of inbreeding depression. *Nature Reviews Genetics*, **10**, 783–796.

Chen, C., **Lu, R.S.**, **Zhu, S.S.**, **Tamaki, I.** & **Qiu, Y.X.** (2017) Population structure and historical demography of *Dipteronia dyeriana* (Sapindaceae), an extremely narrow palaeoendemic plant from China: implications for conservation in a biodiversity hot spot. *Heredity*, **119**, 95–106.

Chen, J., **Glémin, S.** & **Lascoux, M.** (2017) Genetic diversity and the efficacy of purifying selection across plant and animal species. *Molecular Biology and Evolution*, **34**, 1417–1428.

Choi, Y., **Sims, G.E.**, **Murphy, S.**, **Miller, J.R.** & **Chan, A.P.** (2012) Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, **7**, e46688.

Cingolani, P., **Platts, A.**, **Wang, L.L.**, **Coon, M.**, **Nguyen, T.**, **Wang, L.** ***et al.*** (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, **6**, 80–92.

Danecek P, **Auton A**, **Abecasis G**, **Albers CA**, **Banks E**, **DePristo MA**, **Handsaker RE**, **Lunter G**, **Marth GT**, **Sherry ST** etal. 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.

Ding, W.N., **Huang, J.**, **Su, T.**, **Xing, Y.W.** & **Zhou, Z.K.** (2018) An early Oligocene occurrence of the palaeoendemic genus *Dipteronia* (Sapindaceae) from Southwest China. *Review of Palaeobotany and Palynology*, **249**, 16–23.

Do, R., **Balick, D.**, **Li, H.**, **Adzhubei, I.**, **Sunyaev, S.** & **Reich, D.** (2015) No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. *Nature Genetics*, **47**, 126–131.

Excoffier, L., **Dupanloup, I.**, **Huerta-Sánchez, E.**, **Sousa, V.C.** & **Foll, M.** (2013) Robust demographic inference from genomic and SNP data. *PLoS Genetics*, **9**, e1003905.

Eyre-Walker, A. & **Keightley, P.D.** (2007) The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, **8**, 610–618.

Farnsworth, A., **Lunt, D.J.**, **Robinson, S.A.**, **Valdes, P.J.**, **Roberts, W.H.G.**, **Clift, P.D.** ***et al.*** (2019) Past East Asian monsoon evolution controlled by paleogeography, not CO_2. *Science. Advances*, **5**, eaax1679.

Feng, Y., **Comes, H.P.**, **Zhou, X.P.** & **Qiu, Y.X.** (2019) Phylogenomics recovers monophyly and early Tertiary diversification of *Dipteronia* (Sapindaceae). *Molecular Phylogenetics and Evolution*, **130**, 9–17.

Folk, R.A., **Soltis, P.S.**, **Soltis, D.E.** & **Guralnick, R.** (2018) New prospects in the detection and comparative analysis of hybridization in the tree of life. *American Journal of Botany*, **105**, 364–375.

Gargiulo, R., **Aigi, I.**, **Tanel, K.**, **Fay, M.F.** & **Kull, T.** (2018) High genetic diversity in a threatened clonal species, *Cypripedium calceolus* (Orchidaceae), enables long-term stability of the species in different biogeographical regions in Estonia. *Botanical Journal of the Linnean Society*, **186**, 560–571.

Glémin, S. (2003) How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution*, **57**, 2678–2687.

Gómez-Sánchez, D., **Olalde, I.**, **Sastre, N.**, **Enseñat, C.**, **Carrasco, R.**, **Marques-Bonet, T.** ***et al.*** (2018) On the path to extinction: inbreeding and admixture in a declining grey wolf population. *Molecular Ecology*, **27**, 3599–3612.

González-Martínez, S.C., **Ridout, K.** & **Pannell, J.R.** (2017) Range expansion compromises adaptive evolution in an outcrossing plant. *Current Biology*, **27**, 2544–2551.e4.

Grossen, C., **Guillaume, F.**, **Keller, L.F.** & **Croll, D.** (2020) Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nature Communications*, **11**, 1001.

Harrison, S.P., **Yu, G.**, **Takahara, H.** & **Prentice, I.C.** (2001) Diversity of temperate plants in East Asia. *Nature*, **413**, 129–130.

Hartl, D.L. & **Clark, A.G.** (2007) *Principles of population genetics*, 4th edition. Los Angeles: Sinauer Associates.

Hedrick, P.W. & **Garcia-Dorado, A.** (2016) Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology and Evolution*, **31**, 940–952.

Hutchings, J.A. (2015) Thresholds for impaired species recovery. *Proceedings of the Royal Society B: Biological Sciences*, **282**, 20150654.

Jaillon, O., **Aury, J.-M.**, **Noel, B.**, **Policriti, A.**, **Clepet, C.**, **Casagrande, A.** ***et al.*** (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, **449**, 463–467.

Johnson, W.E., **Onorato, D.P.**, **Roelke, M.E.**, **Land, E.D.**, **Cunningham, M.**, **Belden, R.C.** ***et al.*** (2010) Genetic restoration of the Florida panther. *Science*, **329**, 1641–1645.

Kardos, M., **Nietlisbach, P.** & **Hedrick, P.W.** (2018) How should we compare different genomic estimates of the strength of inbreeding depression? *Proceedings of the National Academy of Sciences, USA*, **115**, E2492–E2493.

Keightley, P.D. & **Jackson, B.C.** (2018) Inferring the probability of the derived vs. the ancestral allelic state at a polymorphic site. *Genetics*, **209**, 897–906.

Khan, A., **Patel, K.**, **Shukla, H.**, **Viswanathan, A.**, **van der Valk, T.**, **Borthakur, U.** ***et al.*** (2021) Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. *Proceedings of the National Academy of Sciences*, **118**, e2023018118.

Kimura, M. (1983) *The neutral theory of molecular evolution*. Cambridge, UK: Cambridge University Press.

Korunes, K.L. & **Samuk, K.** (2021) pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, **21**, 1359–1368.

Kyriazis, C.C., **Wayne, R.K.** & **Lohmueller, K.E.** (2021) Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. *Evolution Letters*, **5**, 33–47.

Langmead, B. & **Salzberg, S.L.** (2012) Fast gapped-read alignment with bowtie 2. *Nature Methods*, **9**, 357–359.

Li, H. & **Durbin, R.** (2011) Inference of human population history from individual whole-genome sequences. *Nature*, **475**, 493–496.

Li, Y.Y., **Guan, S.M.**, **Yang, S.Z.**, **Luo, Y.** & **Chen, X.Y.** (2011) Genetic decline and inbreeding depression in an extremely rare tree. *Conservation Genetics*, **13**, 343–347.

Li, Z. & **Barker, M.S.** (2020) Inferring putative ancient whole-genome duplications in the 1000 plants (1KP) initiative: access to gene family phylogenies and age distributions. *GigaScience*, **9**, giaa004.

Lidgard, S. & **Love, A.C.** (2018) Rethinking living fossils. *Bioscience*, **68**, 760–770.

Lin, H.Y., **Li, W.H.**, **Lin, C.F.**, **Wu, H.R.** & **Zhao, Y.P.** (2022) International biological flora: *Ginkgo biloba*. *Journal of Ecology*, **110**, 951–982.

Lindborg, R. & **Eriksson, O.** (2004) Historical landscape connectivity affects present plant species diversity. *Ecology*, **85**, 1840–1845.

Liu, J., **Lindstrom, A.J.**, **Chen, Y.S.**, **Nathan, R.** & **Gong, X.** (2021) Congruence between ocean-dispersal modelling and phylogeography explains recent evolutionary history of *Cycas* species with buoyant seeds. *New Phytologist*, **232**, 1863–1875.

Liu, X. & **Fu, Y.X.** (2020) Stairway plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biology*, **21**, 1–9.

Löve, Á. (1979) IOPB chromosome number reports LXV. *Taxon*, **28**, 5–6.

Lynch, M., **Conery, J.** & **Burger, R.** (1995) Mutation accumulation and the extinction of small populations. *American Naturalist*, **146**, 489–518.

Ma, Q., **Sun, T.**, **Li, S.**, **Wen, J.**, **Zhu, L.**, **Yin, T.** ***et al.*** (2020) The *Acer truncatum* genome provides insights into nervonic acid biosynthesis. *The Plant Journal*, **104**, 662–678.

Manchester, S.R., **Chen, Z.D.**, **Lu, A.M.** & **Uemura, K.** (2009) Eastern Asian endemic seed plant genera and their paleogeographic history throughout the Northern Hemisphere. *Journal of Systematics and Evolution*, **47**, 1–42.

Manichaikul, A., **Mychaleckyj, J.C.**, **Rich, S.S.**, **Daly, K.**, **Sale, M.** & **Chen, W.-M.** (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics*, **26**, 2867–2873.

McClain, A.M. & **Manchester, S.R.** (2001) *Dipteronia* (Sapindaceae) from the Tertiary of North America and implications for the phytogeographic history of the Aceroideae. *American Journal of Botany*, **88**, 1316–1325.

McKey, D., **Elias, M.**, **Pujol, M.E.** & **Duputié, A.** (2010) The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, **186**, 318–332.

Milne, R.I. & **Abbott, R.J.** (2002) The origin and evolution of tertiary relict floras. *Advances in Botanical Research*, **38**, 281–314.

Moritz, C. (1994) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, **9**, 373–375.

Nadachowska-Brzyska, K., **Burri, R.**, **Smeds, L.** & **Ellegren, H.** (2016) PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white *Ficedula* flycatchers. *Molecular Ecology*, **25**, 1058–1072.

Norén, K., **Godoy, E.**, **Dalén, L.**, **Meijer, T.** & **Angerbjörn, A.** (2016) Inbreeding depression in a critically endangered carnivore. *Molecular Ecology*, **25**, 3309–3318.

Ochoa, A. & **Gibbs, H.L.** (2021) Genomic signatures of inbreeding and mutation load in a threatened rattlesnake. *Molecular Ecology*, **30**, 5454–5469.

Oginuma, K.Z., **Gu, Z.**, **Yue, Z.** & **Kondo, K.** (1994) Chromosomes of some woody plants native to Yunnan, China. *Kromosomo*, **73**, 2491–2497.

Ouyang, Z.Q., **Su, W.H.** & **Zhang, G.F.** (2006) Studies on character of seed germination of rare plant *Dipteronia dyeriana*. *Acta Botanica Yunnanica*, **28**, 509–514 [In Chinese with English summary.].

Poon, A. & **Otto, S.P.** (2000) Compensating for our load of mutations: freezing the meltdown of small populations. *Evolution*, **54**, 1467–1479.

Purcell, S., **Neale, B.**, **Todd-Brown, K.**, **Thomas, L.**, **Ferreira, M.A.R.**, **Bender, D.** ***et al.*** (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.

Qiu, Y.-X., **Luo, Y.-P.**, **Comes, H.P.**, **Ouyang, Z.-Q.** & **Fu, C.-X.** (2007) Population genetic diversity and structure of *Dipteronia dyerana* (Sapindaceae), a rare endemic from Yunnan Province, China, with implications for conservation. *Taxon*, **56**, 427–437.

Ramu, P., **Esuma, W.**, **Kawuki, R.**, **Rabbi, I.Y.**, **Egesi, C.**, **Bredeson, J.V.** ***et al.*** (2017) Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. *Nature Genetics*, **49**, 959–963.

Robinson, J.A., **Brown, C.**, **Kim, B.Y.**, **Lohmueller, K.E.** & **Wayne, R.K.** (2018) Purging of strongly deleterious mutations explains long-term persistence and absence of inbreeding depression in Island Foxes. *Current Biology*, **28**, 3487–3494.e4.

Robinson, J.A., **Ortega-Del Vecchyo, D.**, **Fan, Z.**, **Kim, B.Y.**, **Vonholdt, B.M.**, **Marsden, C.D.** ***et al.*** (2016) Genomic flatlining in the endangered Island Fox. *Current Biology*, **26**, 1183–1189.

Robinson, J.A., **Räikkönen, J.**, **Vucetich, L.M.**, **Vucetich, J.A.**, **Peterson, R.O.**, **Lohmueller, K.E.** ***et al.*** (2019) Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Science. Advances*, **5**, eaau0757.

Schiffels, S. & **Durbin, R.** (2014) Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, **46**, 919–925.

Schumann, D., **Raub, T.D.**, **Kopp, R.E.**, **Guerquin-Kern, J.L.**, **Di, W.T.**, **Rouiller, I.** ***et al.*** (2008) Gigantism in unique biogenic magnetite at the Paleocene–Eocene Thermal Maximum. *Proceedings of the National Academy of Sciences, USA*, **105**, 17648–17653.

Sexton, P.F., **Norris, R.D.**, **Wilson, P.A.**, **Pälike, H.**, **Westerhold, T.**, **Röhl, U.** ***et al.*** (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

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

Su, W.H., **Zhang, G.F.** & **Ouyang, Z.Q.** (2006) Community characteristics and conservation strategies of a rare species, *Dipteronia dyerana*. *Acta Botanica Yunnanica*, **28**, 54–58 [In Chinese with English summary.].

Szpiech, Z.A., **Xu, J.**, **Pemberton, T.J.**, **Peng, W.**, **Zöllner, S.**, **Rosenberg, N.A.** ***et al.*** (2013) Long runs of homozygosity are enriched for deleterious variation. *American Journal of Human Genetics*, **93**, 90–102.

Tang, C.Q., **Matsui, T.**, **Ohashi, H.**, **Dong, Y.-F.F.**, **Momohara, A.**, **Herrando-Moraira, S.** ***et al.*** (2018) Identifying long-term stable refugia for relict plant species in East Asia. *Nature Communications*, **9**, 4488.

Tang, C.Q., **Yang, Y.**, **Ohsawa, M.**, **Momohara, A.**, **Hara, M.**, **Cheng, S.** ***et al.*** (2011) Population structure of relict *Metasequoia glyptostroboides* and its habitat fragmentation and degradation in south-central China. *Biological Conservation*, **144**, 279–289.

Tang, C.Q., **Yang, Y.**, **Ohsawa, M.**, **Yi, S.R.**, **Momohara, A.**, **Su, W.H.** ***et al.*** (2012) Evidence for the persistence of wild *Ginkgo biloba* (Ginkgoaceae) populations in the Dalou mountains, southwestern China. *American Journal of Botany*, **99**, 1408–1414.

Terhorst, J., **Kamm, J.A.** & **Song, Y.S.** (2016) Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, **49**, 303–309.

Tian, D., **Patton, A.H.** & **TurnerBJ, M.C.H.** (2022) Severe inbreeding, increased mutation load and gene loss-of-function in the critically endangered Devils Hole pupfish. *Proceedings of the Royal Society B*, **289**, 20221561.

van Oosterhout, C. (2020) Mutation load is the spectre of species conservation. *Nature Ecology and Evolution*, **4**, 1004–1006.

Vargas, P., **Jiménez-Mejías, P.** & **Fernández-Mazuecos, M.** (2020) 'Endangered living fossils' (ELFs): long-term survivors through periods of dramatic climate change. *Environmental and Experimental Botany*, **170**, 103892.

Vaser, R., **Adusumalli, S.**, **Leng, S.N.**, **Sikic, M.** & **Ng, P.C.** (2016) SIFT missense predictions for genomes. *Nature Protocols*, **11**, 1–9.

Welch, J.J., **Eyre-Walker, A.** & **Waxman, D.** (2008) Divergence and polymorphism under the nearly neutral theory of molecular evolution. *Journal of Molecular Evolution*, **67**, 418–426.

Werth, A.J. & **Shear, W.A.** (2014) The evolutionary truth about living fossils. *American Scientist*, **102**, 434–443.

Whitlock, M.C. & **Bürger, R.** (2009) Fixation of new mutations in small populations. In: Ferrière, R., Dieckmann, U. & Couvet, D. (Eds.) *Evolutionary conservation biology*. Cambridge, UK: Cambridge University Press, pp. 155–170.

Wolfe, J.A. & **Tanai, T.** (1987) Systematics, phylogeny, and distribution of *acer* (maples) in the Cenozoic of western North America. *Journal of the Faculty of Science, Hokkaido University, Ser. IV*, **22**, 1–246.

Wright, S. (1931) Evolution in mendelian populations. *Genetics*, **16**, 97–159.

Xu, T.Z., **Chen, Y.S.**, **de Jong, P.C.**, **Oterdoom, H.J.** & **Chang, C.S.** (2008) Aceraceae. In: Wu, Z.Y., Raven, P.H. & Hong, D.Y. (Eds.) *Flora of China*. Beijing, China: Science Press; St. Louis, USA: Missouri Botanical Garden Press, pp. 515–553.

Xue, Y., **Prado-Martinez, J.**, **Sudmant, P.H.**, **Narasimhan, V.**, **Ayub, Q.**, **Szpak, M.** ***et al.*** (2015) Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science*, **348**, 242–245.

Yang, J., **Lee, S.H.**, **Goddard, M.E.** & **Visscher, P.M.** (2011) GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, **88**, 76–82.

Yang, J., **Wariss, H.M.**, **Tao, L.**, **Zhang, R.**, **Yun, Q.**, **Hollingsworth, P.** ***et al.*** (2019) De novo genome assembly of the endangered *Acer yangbiense*, a plant species with extremely small populations endemic to Yunnan Province, China. *GigaScience*, **8**, 1–10.

Yang, Y., **Ma, T.**, **Wang, Z.**, **Lu, Z.**, **Li, Y.**, **Fu, C.** ***et al.*** (2018) Genomic effects of population collapse in a critically endangered ironwood tree *Ostrya rehderiana*. *Nature Communications*, **9**, 1–9.

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.

Yates, M.C., **Bowles, E.** & **Fraser, D.J.** (2019) Small population size and low genomic diversity have no effect on fitness in experimental translocations of a wild fish. *Proceedings of the Royal Society B: Biological Sciences*, **286**, 20191989.

Young, A.G., **Boyle, T.** & **Brown, T.** (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, **11**, 413–419.

Zhang, H.H. (2000) *The threatened wild plants in Guizhou Province*. Beijing, China: China Forest Press.

Zhao, Y.-P., **Fan, G.**, **Yin, P.-P.**, **Sun, S.**, **Li, N.**, **Hong, X.** ***et al.*** (2019) Resequencing 545 *Ginkgo* genomes across the world reveals the evolutionary history of the living fossil. *Nature Communications*, **10**, 1–10.

Zheng, B., **Xu, Q.** & **Shen, Y.** (2002) The relationship between climate change and Quaternary glacial cycles on the Qinghai–Tibetan Plateau: review and speculation. *Quaternary International*, **97–98**, 93–101.

Zhu, S., **Chen, J.**, **Zhao, J.**, **Comes, H.P.**, **Li, P.**, **Fu, C.** ***et al.*** (2020) Genomic insights on the contribution of balancing selection and local adaptation to the long-term survival of a widespread living fossil tree, *Cercidiphyllum japonicum*. *New Phytologist*, **228**, 1674–1689.