Molecular Ecology Resources

WILEY

MOLECULAR ECOLOGY RESOURCES

RESOURCE ARTICLE

Chromosomal-Level Genome Suggests Adaptive Constraints Leading to the Historical Population Decline in an Extremely Endangered Plant

Shao Shao[1] | Yulong Li[1,2] | Xiao Feng[1] | Chuanfeng Jin[1] | Min Liu[1] | Ranran Zhu[1] | Miles E. Tracy[1] | Zixiao Guo[1] | Ziwen He[1] | Suhua Shi[1] | Shaohua Xu[1,2]

[1]State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Stress Biology, Innovation Center for Evolutionary Synthetic Biology, School of Life Sciences, Sun Yat-sen University, Guangzhou, China | [2]School of Ecology, Sun Yat-sen University, Shenzhen, China

Correspondence: Suhua Shi (lssssh@mail.sysu.edu.cn) | Shaohua Xu (xushh27@mail.sysu.edu.cn)

Received: 15 January 2024 | **Revised:** 14 October 2024 | **Accepted:** 28 October 2024

Handling Editor: Alison Gonçalves Nazareno

Funding: This work was supported by the National Natural Science Foundation of China (32330005, 32100176 and 32370227), the Guangdong Basic and Applied Basic Research Foundation (2022A1515010962 and 2023B1515020083) and the China Postdoctoral Science Foundation (2024T170174).

Keywords: conservation genomics | evolutionary genomics | non-dormant seeds | *Pellacalyx yunnanensis* | plant species with extremely small populations

ABSTRACT

Increased human activity and climate change have significantly impacted wild habitats and increased the number of endangered species. Exploring evolutionary history and predicting adaptive potential using genomic data will facilitate species conservation and biodiversity recovery. Here, we examined the genome evolution of a critically endangered tree *Pellacalyx yunnanensis*, a plant species with extremely small populations (PSESP) that is narrowly distributed in Xishuangbanna, China. The species has neared extinction due to economic exploitation in recent decades. We assembled a chromosome-level genome of 334 Mb, with the N50 length of 20.5 Mb. Using the genome, we discovered that *P. yunnanensis* has undergone several population size reductions, leading to excess deleterious mutations. The species may possess low adaptive potential due to reduced genetic diversity and the loss of stress-responsive genes. We estimate that *P. yunnanensis* is the basal species of its genus and diverged from its relatives during global cooling, suggesting it was stranded in unsuitable environments during periods of dramatic climate change. In particular, the loss of seed dormancy leads to germination under unfavourable conditions and reproduction challenges. This dormancy loss may have occurred through genetic changes that suppress ABA signalling and the loss of genes involved in seed maturation. The high-quality genome has also enabled us to reveal phenotypic trait evolution in Rhizophoraceae and identify divergent adaptation to intertidal and inland habitats. In summary, our study elucidates mechanisms underlying the decline and evaluates the adaptive potential of *P. yunnanensis* to future climate change, informing future conservation efforts.

1 | Introduction

Due to the joint impact of climate change and human activities, the rate of biological extinctions worldwide is accelerating rapidly, posing severe challenges to global biodiversity (Díaz et al. 2019; Pimm et al. 2014; Urban 2015). Plant species with extremely small populations (PSESP) constitute the most urgent and valuable group among endangered species. PSESPs are characterised by exceedingly small population sizes and individual numbers, narrow habitats, exposure to severe human disturbance and a high risk of extinction (Ma et al. 2013; Sun, Yang, and Dao 2019; Yang et al. 2020). They often face difficulties in environmental adaptation and reproduction and are easily affected by human activities.

The protection of PSESPs holds significant biological importance. As the most vulnerable biological resources at risk of being lost, PSESPs are typically endemic to specific regions. Their unique ecological, economic and cultural value are at risk of disappearing with species extinction. Moreover, PSESPs can potentially reveal the impact of geological history and climate change on plant communities. In recent years, genomic data have shown great potential in guiding endangered species conservation (Paez et al. 2022). Genomic studies of PSESPs have indicated that genetic drift and inbreeding increases in small populations over extended periods lead to reduced genetic diversity, widespread homozygosity and an accumulation of genetic load (Z. Chen et al. 2020; Ma et al. 2021; Yang et al. 2022; Yang et al. 2018). However, comprehensive genomic studies on the evolutionary history and adaptive potential of PSESPs remain scarce. Genomic analysis can elucidate the mechanisms of environmental adaptation defects and reproductive challenges in PSESPs.

In this study, we apply genomic data to protect *Pellacalyx yunnanensis*, a PSESP restrictedly distributed in the forests of Xishuangbanna, Yunnan, China. The species has been classified as Endangered (EN) on the IUCN Red List and the China Biodiversity Red List (Qin et al. 2017; World Conservation Monitoring Centre 1998; Wu and Raven 1994). Over the years, the large-scale cultivation of economic crops has led to destructive logging of *P. yunnanensis*, with many original distribution areas now devoid of surviving individuals (Su et al. 2005). For instance, *P. yunnanensis* in Jingpiao has been replaced by *Wurfbainia villosa*, and its populations in Dakaqing have been converted to rubber plantations. Consequently, only two natural populations and one ex situ conservation group with 33 individuals survived. Furthermore, *P. yunnanensis* has experienced a significant bottleneck effect due to human interference, exhibiting large inter-population and low intra-population diversity (Su et al. 2005). This places *P. yunnanensis* at the highest threat level among the National Key Protected Wild Plants in Xishuangbanna (Xu and Tao 1987). Notably, this species could be regarded as a representative endangered plant of Xishuangbanna, where 15.68% of China's endangered plants are found on just 0.2% of its land area (Wen and Song 2005).

Like many other endangered endemic species, *P. yunnanensis* does not share habitats with the other seven species of its genus (Figure 1a). The species occupies tropical rainforests in southern Yunnan, which are at the northern edge of the tropics and experience relatively low annual temperatures and rainfall (Zhu 2022). In contrast, the other *Pellacalyx* species inhabit Southeast Asian tropical rainforests in Thailand, Myanmar, Malaysia and the Philippines, where temperatures and humidity are higher (Juncosa and Tomlinson 1988a). A genomic study of *P. yunnanensis* would enhance our understanding of the impact of geological events and climate change on forming PSESPs.

Diminished reproductive capacity is a critical factor in the endangerment of PSESPs and warrants research into the underlying mechanisms. *Pellacalyx yunnanensis* seed exhibits maladaptive germination in current environmental conditions. The seeds are non-dormant and attempt to germinate immediately after maturation in winter when typical temperatures are below 19°C, despite requiring temperatures above 25°C for successful sprouting (Ma et al. 1988). Even when temperatures are suitable from April to May the following year, the humidity remains too low for germination. Additionally, long-term seed storage leads to a rapid decline in germination rates. The difficulties in natural reproduction, compounded by human activities, have precipitated a sharp decrease in the habitat area and population size of *P. yunnanensis*.

The species is recognised as one of the closest inland relatives of the largest mangrove clade Rhizophoreae and shares several phenotypic traits with its mangrove relatives (Juncosa and Tomlinson 1988b; Schwarzbach and Ricklefs 2000; Setoguchi, Kosuge, and Tobe 1999; Shi et al. 2002; Xu et al. 2017). Mangroves are woody plants that evolved from inland ancestors to inhabit tropical land–sea interface, developing specialised traits like aerial roots and non-dormant viviparous seeds to adapt to the tropical intertidal zones (Duke 1992; Shi et al. 2005; Tomlinson 2016; Xu et al. 2017). *Pellacalyx yunnanensis* also has aerial roots and non-dormant seeds similar to its Rhizophoreae mangrove relatives. These similar traits may have helped it initially adapt to tropical rainforests, but were not conducive to its adaptation to the mountains at the northern edge of the tropics where it now lives (Ma et al. 1988). The *P. yunnanensis* genome may help identify adaptation mechanisms of both inland and intertidal environments.

To guide PSESP conservation efforts, we de novo assembled a chromosome-level genome of *P. yunnanensis*, explored the factors affecting this species' origin and narrow distribution and assessed its genome vulnerability. Additionally, we examined mechanisms underlying seed non-dormancy and compared genomic features with those of mangroves to identify adaptive signatures for either inland or intertidal environments. Our findings suggest that genomic sequences can help reveal the mechanisms underlying population decline and evaluate the adaptive potential of PSESPs, thereby informing future conservation efforts.

2 | Materials and Methods

2.1 | Genome Sequencing

We extract genomic DNA from fresh leaves of *Pellacalyx yunnanensis* for whole genome sequencing. *Pellacalyx yunnanensis* leaves were collected from an adult individual in Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Yunnan, China. Fresh leaves were quick-frozen in liquid nitrogen and stored in a freezer at −80°C. Two grams of leaves were ground in liquid nitrogen, and genomic DNA was extracted using the CTAB method (Doyle and Doyle 1987).

Then, genomic DNA was used to construct a sequencing library and sequenced using Single Molecule Real Time (SMRT) sequencing technology. To construct the PacBio library for assembly, at least 10 μg genomic DNA was sheared into fragments of approximately 20 kb, and fragments < 7 kb were filtered out using BluePippin (Sage Science, MA, USA). The filtered DNA was transformed into a proprietary SMRTbell library using the PacBio DNA Template Preparation Kit. SMRT sequencing was performed using HiFi Bundle (V2) sequencing reagents and SMRT Cell (8 M) on the PacBio Sequel II sequencing platform.

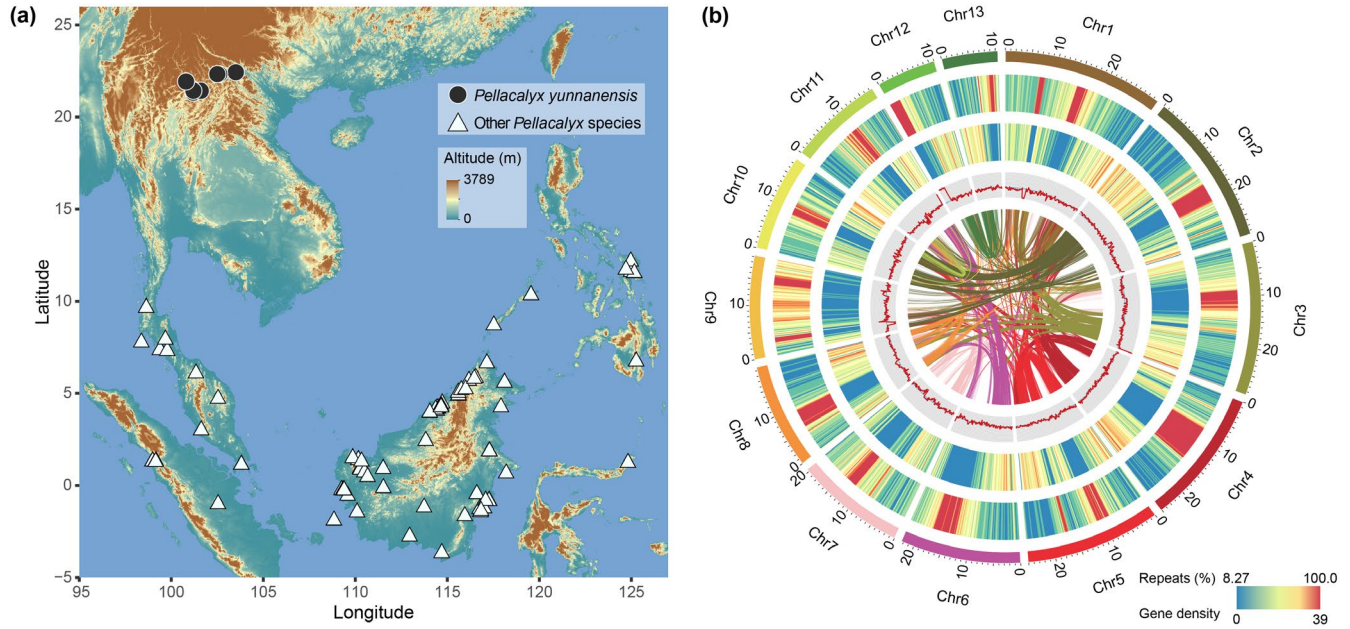

FIGURE 1 | Distribution and genomic characteristics of *P. yunnanensis*. (a) Natural distribution of *Pellacalyx* species. Each point represents a specimen, and the data come from the National Specimen Information Infrastructure (NSII) and Global Biodiversity Information Facility (GBIF) (http://www.nsii.cn; https://www.gbif.org). The digital elevation model comes from GMTED2010 (Danielson and Gesch 2011). (b) Genome features of *P. yunnanensis*. The coloured bars on the outer track demarcate the 13 chromosomes. Circular tracks represent the percentage of repeats, gene density and GC content from outside to inside (200 kb windows). Each linking line in the circle's centre connects a duplicated gene pair.

In addition to the SMRT sequencing, we obtained DNA paired-end (PE) short reads for assembly calibration and sequencing error correction and RNA-sequencing data for gene prediction. We use an ultrasonic processing device to prepare DNA short-insert PE libraries. The 300 bp PE library was built using the Illumina library preparation kit and was sequenced on the Illumina X-Ten platform (San Diego, CA, USA). Raw reads were trimmed using Trimmomatic (v.0.35) (Bolger, Lohse, and Usadel 2014). Illumina sequence adaptors were removed, low-quality bases from the start or end of raw reads were trimmed and reads were scanned using a four-bp sliding window and trimmed when the average quality per base dropped below 15.

The total RNA used to annotate expressed proteins was extracted from 100 mg leaves using TRIzol reagent (Invitrogen). The RNA-Seq library was constructed using reagents provided by the NEBNext Ultra RNA Library Prep Kit (Illumina, USA). The mRNA molecules were first fragmented and copied into first-strand cDNA using reverse transcriptase and random hexamer primers, and second-strand cDNA was synthesised using DNA polymerase I and RNase H. The cDNA fragments were then ligated with adapters and enriched by polymerase chain reaction (PCR). The cDNA molecules were sequenced on the Illumina X-TEN platform and the raw reads were trimmed using Trimmomatic.

2.2 | De Novo Genome Assembly

Genome assembly was first obtained using PacBio HiFi reads, then corrected using DNA short reads, and finally, its quality was evaluated. De novo genome assembly with 390 Gb (~1161×) PacBio HiFi reads was conducted using HiFiasm (v0.14) (Cheng et al. 2021) with the parameters −t32 and −f39, which could speed up assembly. Then, to reduce the assembly error rate, 53 Gb (~158×) DNA short reads were aligned to the assembled contigs by BWA (v0.7.17) (Li and Durbin 2009), and several rounds of iterative correction were performed using Pilon (v1.2) (Walker et al. 2014) with default parameters.

Assembly integrity was assessed using three methods: (1) The assembled genome size was compared with that estimated one using *k*-mer method. Based on the DNA short read library, the *k*-mer depth distribution ($K=17$) was obtained, and the genome size was estimated by GenomeScope (v1.0.0) (Vurture et al. 2017). (2) The existence of the conserved single-copy orthologues among eudicots was used to evaluate assembly completeness, and these sequences were searched against the genome assembly using the BUSCO software (v5.0.0) (Manni et al. 2021; Seppey, Manni, and Zdobnov 2019). (3) The DNA short reads and RNA-Seq reads were aligned to the genome assembly using BWA and HISAT2 (v2.2.1) (Kim et al. 2019). Higher mapping rates indicate higher assembly completeness.

2.3 | Repeat Sequence and Protein-Coding Gene Annotation

To prevent interference with protein-coding gene prediction, repetitive sequences were initially identified using RepeatModeler (v2.02a) (Flynn et al. 2020) and RepeatMasker (v4.1.2-p1) (Tarailo-Graovac and Chen 2009). The most abundant repetitive sequences in plant genomes, long terminal repeat retrotransposons (LTRs), are also identified using LTR_retriever (v2.8.7) (Ou and Jiang 2018).

After repeat sequences were identified and masked, protein-coding genes were predicted using a combination of de novo, homology-based and transcriptome-based predictions. De novo gene predictions identify genes according to gene features and have high sensitivity but relatively low accuracy. Homology-based and transcriptome-based predictions have high accuracy but relatively lower sensitivity. In practice, the three methods are usually used in combination. De novo gene predictions were conducted using Augustus (v2.5.5) (Stanke et al. 2006) and GeneMark (v4.32) (Lomsadze et al. 2005). Homology-based predictions were performed using EXONERATE (v2.2.0) (Slater and Birney 2005). Protein sequences of *Arabidopsis thaliana*, *Linum usitatissimum*, *Manihot esculenta*, *Medicago truncatula*, *Mimulus guttatus*, *Populus trichocarpa*, *Rhizophora apiculata*, *Ricinus communis*, *Theobroma cacao* and *Vitis vinifera* were searched against the genome assembly for homology-based predictions (Bredeson et al. 2016; Chan et al. 2010; He et al. 2022; Hellsten et al. 2013; Lamesch et al. 2012; Motamayor et al. 2013; Tang et al. 2014; The French–Italian Public Consortium for Grapevine Genome Characterization 2007; Tuskan et al. 2006; Z. Wang et al. 2012). Transcriptome-based predictions were conducted by mapping RNA-Seq data (8.57 Gb, ~26×) to the genome assembly using HISAT2 and StringTie (v1.3.5) (Kovaka et al. 2019). EVidenceModeler (EVM) (Haas et al. 2008) integrates the gene models obtained by the three methods into a non-redundant gene set.

Functional annotations of gene sets were obtained by searching against NCBI non-redundant and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases using BLAST (v2.2.6; −e 1e-5) (Camacho et al. 2009); protein domains and Gene Ontology (GO) annotations were obtained by searching against the InterPro databases by InterProScan (Jones et al. 2014). Genomic features such as chromosome length, TE content, gene density, GC content and duplicated gene pairs were counted in a 200 kb window and visualised by Circos (v0.69-9) (Krzywinski et al. 2009).

2.4 | Phylogenetic Analysis and Divergence Time Estimation

The phylogenetic tree reconstruction and divergence time estimation were conducted using a two-step method (He et al. 2022). The first step is to construct a main tree based on the protein sequences of *P. yunnanensis* and nine high-quality genomes: *Rhizophora apiculata*, *Bruguiera gymnorhiza*, *Carallia pectinifolia*, *Erythroxylum novogranatense, Populus trichocarpa*, *Ricinus communis, Arabidopsis thaliana*, *Vitis vinifera* and *Oryza sativa* (He et al. 2022; Ouyang et al. 2007; Wang et al. 2023). The 996 single-copy orthologous gene families were identified by OrthoFinder (v2.3.11; -I 1.5) (Emms and Kelly 2019). Protein sequences for each gene family were aligned by MAFFT (v7.429; --maxiterate 1000 -localpair) (Katoh and Standley 2013) and converted into the corresponding nucleotide sequence by PAL2NAL (v14) (Suyama, Torrents, and Bork 2006). After removing low-quality alignment regions using Gblocks (v0.91; −b5 = *n*) (Castresana 2000), the single-copy orthologous sequences of each species were concatenated end to end into a super-sequence. These sequences were used to construct the phylogenetic tree using the GTR + G model repeated 1000 times by RAxML-NG (v1.0.2) (Kozlov et al. 2019).

These sequences were also used to calculate divergence times using the approximate likelihood method of the MCMCTree program in PAML (v4.9j) (Yang 2007), which was repeated 10,000,000 times to check convergence. Four reliable fossils were used to correct divergence times: (1) the separation time of monocots and eudicots < 247 million years ago (Ma); (2) the root of core eudicots < 125 Ma; (3) the most recent ancestor of Rhizophoreae mangroves from the early Eocene (56–47.8 Ma); and (4) the earliest ancestor of the *Rhizophora* genus from the late Eocene (37.8–33.9 Ma) (Graham 2006; Morris et al. 2018; Muller 1981).

In the second step, we added several DNA fragments from seven species: *Pellacalyx axillaris*, *Pellacalyx lobbii, Pellacalyx saccardianus*, *Crossostylis biflora*, *Gynotroches axillaris*, *Cassipourea ceylanica* and *Erythroxylum novogranatense* (Schwarzbach and Ricklefs 2000). DNA fragments include *rbcL*, *atp2-rbcL* spacer, *trnL-trnF* spacer and nuclear ribosomal DNA sequences (ITS1, ITS2, 5.8S and partial 26S). The published sequences were downloaded from NCBI, and the corresponding sequences of *P. yunnanensis*, *Ca. pectinifolia* and Rhizophoreae mangroves were searched in the genome and chloroplast genome using BLAT (v36 × 4) (Kent 2002). All nucleotide sequences were aligned using MAFFT and filtered with Gblocks. The four fragment sequences were concatenated for tree construction using the same method as in the first step. Time nodes were inferred using MCMCTree and corrected using the times of the main tree. Two hypermetric trees were merged into one final time tree.

2.5 | Whole-Genome Duplication Identification

Ancient whole-genome duplication (WGD) events are prevalent in plant genome evolution. After experiencing WGD, the plant species became polyploid but eventually returned to diploidy following extensive gene loss and genome rearrangements. The traces of ancient WGDs on the genome are the collinear blocks that contain homologous genes arranged in the same order. To identify ancient WGD events, we aligned all protein sequences using BLASTp (v2.11.0; *e*-value $< 10^{-10}$; identity ≥ 40%). Duplicate genes located on the same chromosome and separated by no more than five genes were defined as tandem duplications. After excluding tandem duplications, we defined two genomic regions with at least five pairs of homologous genes arranged in the same order as collinear blocks using *MCScanX* (Y. Wang et al. 2012).

These intraspecific collinear blocks also contain traces of more ancient WGD events (Van de Peer, Mizrachi, and Marchal 2017). To focus on the recent lineage-specific WGD event shared by Rhizophoreae mangroves and inland plants (*P. yunnanensis* and *Ca. pectinifolia*), we filtered out gene pairs resulting from ancient WGD events using a synonymous substitution rate (Ks) cut-off (Xu et al. 2017). Sequence alignment of gene pairs was performed using MAFFT and PAL2NAL, and Ks was calculated by KaKs_Calculator (v2.0; -m YN) (Wang et al. 2010). According to the Ks distribution, collinear blocks with median Ks < 1.25 were

retained (Figure S12). Gene pairs with Ks still greater than 1.5 in these collinear blocks may not belong to this recent WGD event and were removed. In addition, to exclude very recent segmental duplications, gene pairs with Ks < 0.1 within these collinear blocks were also removed. We used the same Ks interval to filter interspecific collinear blocks. WGD-duplicated families with more than half of the mangroves but none of the inland plants represented are retained as unique to mangroves; otherwise, they are defined as unique to inland plants. We used the 'clusterProfiler' package (Yu et al. 2012) for GO analyses.

The WGD event could be dated by the genetic distance between paralogs in collinear blocks. Given the branch length (L) between the WGD event and the node closest to it (node_c), the time from WGD to node_c can be estimated by L/μ, where μ is the average mutation rate from node_c to the present.

2.6 | Heterozygosity Estimation and Population History Inference

To assess genomic heterozygosity, clean DNA short reads were aligned to the reference genome using BWA (v0.7.17). SAMtools (v1.16) (Li et al. 2009) was used to remove incorrect pairings and single-end mappings in alignments, and Picard (v2.20.2; AddOrReplaceReadGroups; https://github.com/broadinstitute/picard) was used to assign read group information including sample identity, library and lane. Sequences were realigned using GATK-3.7 (Depristo et al. 2011; McKenna et al. 2010) to improve calling accuracy, and sites with depth ≥ 20× were retained as effective sites. To minimise the impact of sequencing and mapping biases, we removed low-quality sites based on the following criteria: (1) quality score < 30 and quality corrected by depth (QualByDepth, QD) < 10; (2) probability of strand bias (FisherStrand, FS) > 10; (3) difference between the positions of the reference and alternate alleles in the reads (ReadPosRankSum) < −8. The sites with a minor variant allele frequency ≥ 0.15 were considered heterozygous, and the genomic heterozygosity is estimated as the number of heterozygous sites divided by the number of effective sites. The effects of mutations on protein function were annotated using SnpEff (v4.3) (Cingolani et al. 2012). According to the destructiveness of mutations to protein, they are classified as 'low,' 'tolerable' and 'deleterious' (https://pcingola.github.io/SnpEff/snpeff/inputoutput/). 'Low' variants, such as synonymous mutations, are mostly harmless or unlikely to change protein behaviour. 'Tolerant' variants, such as non-synonymous mutations and codon insertions/deletions, are non-disruptive and may change protein effectiveness. 'Deleterious' variants have destructive effects on the protein, such as start codon deletion and frameshift mutations. Population history was inferred using PSMC (v0.6.5) (Li and Durbin 2011) with the parameters "−p 4 + 25 * 2 + 4 + 6", repeated 100 times. Mutation rate estimation methods refer to Xu et al. (2017), and the generation time was set at 20 years.

2.7 | Disease Resistance Gene (*R*-Gene) Classification

Hidden Markov models are used to predict the domains of protein sequences. The NBS (PF00931), TIR (PF01582), RPW8 (PF05659), Coiled-coil (PF18052) and multiple LRR domains (PF00560, PF07723, PF07725, PF12799, PF13306, PF13516, PF13504, PF13855 and PF08263) of disease resistance genes were identified using HMMER (v3.2.1; −E 1e-4; http://hmmer.org) and the Pfam database (Mistry et al. 2021). After *R*-genes were identified, we used TBtools (v2.016) (Chen et al. 2023) to visualise their distribution on chromosomes.

We constructed a phylogenetic tree of disease resistance genes in *P. yunnanensis* and *Ca. pectinifolia* to examine their evolution. The NBS domain protein sequences of all disease resistance genes and the flanking 20 amino acids were aligned using MAFFT. After excluding short and highly differentiated sequences, the alignment results were converted into nucleic acid sequences using PAL2NAL. The phylogeny was constructed using the GTR + G model with 1000 iterations by RAxML-NG.

2.8 | Identification of Candidate Genes Involved in Seed Dormancy and Aerial Root Development

We conducted gene family clustering and identified seed dormancy-related genes that belong to the same gene family as known seed genes in *Arabidopsis thaliana*. Proteins from 11 core eudicots, *P. yunnanensis*, *Ca. pectinifolia*, *R. apiculata*, *Ce. tagal*, *K. candel*, *B. gymnorhiza*, *Salix purpurea* (Zhou et al. 2020), *Populus trichocarpa*, *Ricinus communis*, *A. thaliana* and *V. vinifera*, were used for gene family clustering. All protein sequences were aligned using DIAMOND (v0.9.24.125; −e 1e-10; identity ≥ 40%) (Buchfink, Xie, and Huson 2014) and clustered by OrthoFinder. For subsequent analysis, we retained families with more than four genes from at least two species.

Genes related to seed dormancy and aerial root development were collected from The Arabidopsis Information Resource (TAIR) database (Lamesch et al. 2012). Gene families comprising these *A. thaliana* genes were considered candidates. Gene families with more than twice the number of genes in *P. yunnanensis* compared to other inland plants were marked as expanded and those with half the number in *P. yunnanensis* as reduced. The gene number reductions were further confirmed by identifying gene loss events. BLAT was used to examine whether the genes lost in *P. yunnanensis* have homologous segments on the genome. If so, homologous fragments were checked for premature termination and frameshift mutations using Genewise (2.4.1) (Birney, Clamp, and Durbin 2004).

3 | Results

3.1 | De Novo Assembly and Annotation of *Pellacalyx yunnanensis* Genome

We obtained 390 Gb of PacBio HiFi raw reads and 53.1 Gb of DNA short reads for *Pellacalyx yunnanensis* genome assembly (Table S1; Figure S1). The total length of the assembled genome is 334.07 Mb, with an N50 length of 20.5 Mb (Tables 1 and S2; Figure S2). The karyotype of *P. yunnanensis* is $2n = 26$, and the total length of the top 13 longest scaffolds accounted for 82.76%

TABLE 1 | Assembly and annotation features of *P. yunnanensis* genome.

Genome features	*P. yunnanensis*
Total assembly length (Mb)	334.07
Number of scaffolds	986
N50 (Mb)	20.49
BUSCO completeness (%)	97.50
GC content (%)	35.92
Repeats (%)	47.15
Number of protein-coding genes	25,930
Annotated protein-coding genes (%)	99.18

of the genome, suggesting the assembly reached the chromosomal level (Figure 1b) (Oginuma and Tobe 1997). The complete BUSCOs of the genome accounted for 97.5% and 95.0% of DNA short reads and 97.2% of RNA-seq reads could be appropriately mapped to the assembly (Table S3), suggesting a high quality and completeness.

Based on homology search methods and de novo prediction, we found that the *P. yunnanensis* genome contains 47.15% repetitive sequences, of which 86.75% are transposable elements (TEs) (Table S4). Long terminal repeats (LTRs) are the most abundant TEs, and *Gypsy* and *Copia* elements account for 12.03% and 4.83% of the genome, respectively. Most complete LTRs appeared in the genome in the past one million years, with a peak in the past 200,000 years (Figure S3). The repeat sequence content in the *P. yunnanensis* genome is similar to its inland relative *Carallia pectinifolia* and higher than in the Rhizophoreae mangroves that experienced TE reduction. After the repetitive sequences were masked, we predicted 25,930 protein-coding genes. The average gene length is 3177 bp, the average exon number is 5.55 and the average intron length is 416 bp (Table S5). By searching against the InterPro, GO, KEGG, and NR databases, we assigned candidate functions to 99.18% of the genes (Table S6).

3.2 | Evolutionary History of *P. yunnanensis*

To explore the factors underlying population decline, we inferred the origin of *P. yunnanensis* using the genomic data. We first reconstructed a phylogenetic tree using 996 high-confidence single-copy orthologues from 10 plant genomes (Figure S4; Table S7). We calculated divergence times using the MCMCTree program and corrected them using four fossils (Yang 2007). We then added DNA sequences from seven other species, including three *Pellacalyx* taxa, to increase the resolution (Figure S5; Table S7) (Schwarzbach and Ricklefs 2000). Finally, we calculated the divergence time of the subtree and merged it with the main tree.

We found that *P. yunnanensis* clusters with inland plants, forming the Gynotrocheae tribe, and diverged from Rhizophoreae mangrove ancestors 50.6 million years ago (Ma) (Figure 2a; Figure S6). This rejects the hypothesis that *P. yunnanensis* originated from mangrove ancestors (Ma et al. 1988; Zhu 2022). The results also suggest that *P. yunnanensis* was the basal species of the genus and diverged from the ancestor of the other taxa 12.7 Ma. This coincides with the global cooling period (until 10 Ma) following the Mid-Miocene Climatic Optimum (MMCO; 17 Ma – 15 Ma) (Zachos et al. 2001). The continuous decline in temperatures during this period led to the reconstruction of Antarctica's major ice sheets and the emergence of continental glaciers, while southern Yunnan, the current *P. yunnanensis* habitat, was not directly affected by glaciers. The alignment of the phylogenetic inference with geographical events suggests that the latter may have influenced the origin and evolution of *P. yunnanensis*.

Our phylogeny also helps reconstruct the evolutionary history of Rhizophoraceae, whose time of origin and separation into three tribes remain unresolved (Schwarzbach and Ricklefs 2000). Rhizophoraceae diverged from their sister family Erythroxylaceae 80.9 Ma. Macarisieae, the basal Rhizophoraceae tribe, diverged 76.7 Ma shortly after the Rhizophoraceae origin. Adding information from inter- and intraspecific synonymous substitution rates (Ks), we found that the other two tribes, Rhizophoreae and Gynotrocheae, share a whole-genome duplication (WGD) event (69 Ma – 66 Ma) near the K-Pg boundary (Figure 2a,b) (Renne et al. 2013; Xu et al. 2017). They diverged 50.6 Ma after experiencing a series of early Eocene hyperthermal events, including the Palaeocene–Eocene Thermal Maximum (PETM). Rhizophoreae mangroves began invading intertidal zones between 50.6 and 47.8 Ma (Zachos, Dickens, and Zeebe 2008). The basal inland Gynotrocheae genus *Carallia* originated during the middle Eocene (41.8 Ma), while *Crossostylis*, *Gynotroches* and *Pellacalyx* continued to differentiate until the late Oligocene (26.82 and 23.81 Ma).

Since Rhizophoreae mangroves have 18 chromosomes, whereas the Gynotrocheae plants have 12–14 chromosomes, this suggests dramatic chromosomal rearrangements after species divergence (Oginuma and Tobe 1997; Senakun 2018). We identified homologous chromosomal regions by finding orthologous collinear blocks in the *P. yunnanensis*, *Ca. pectinifolia* and Rhizophoreae mangrove genomes. The karyotype of Rhizophoreae mangroves is relatively stable, while Gynotrocheae has undergone many chromosomal fissions and fusions (Figure 2c). The difference in chromosomal numbers should result from chromosome fission in mangroves or chromosome fusion in Gynotrocheae plants. To distinguish the two possible scenarios, we examined intra-chromosomal collinearity blocks. Collinearity blocks are produced by WGDs and should be distributed on different chromosomes. If they were observed in genomic regions of the same chromosome, then chromosome fusions should have occurred. We found 385 pairs of collinear genes on chromosomes 6, 7, 10 and 11 of *Ca. pectinifolia*, while *P. yunnanensis* had 306 pairs distributed along chromosomes 1, 3, 5 and 9 (Figure S7). The number of intra-chromosomal collinear blocks is significantly greater than in *R. apiculata* and *B. gymnorhiza* (33 and 35 pairs), suggesting chromosome fusions in *P. yunnanensis* and *Ca. pectinifolia*.

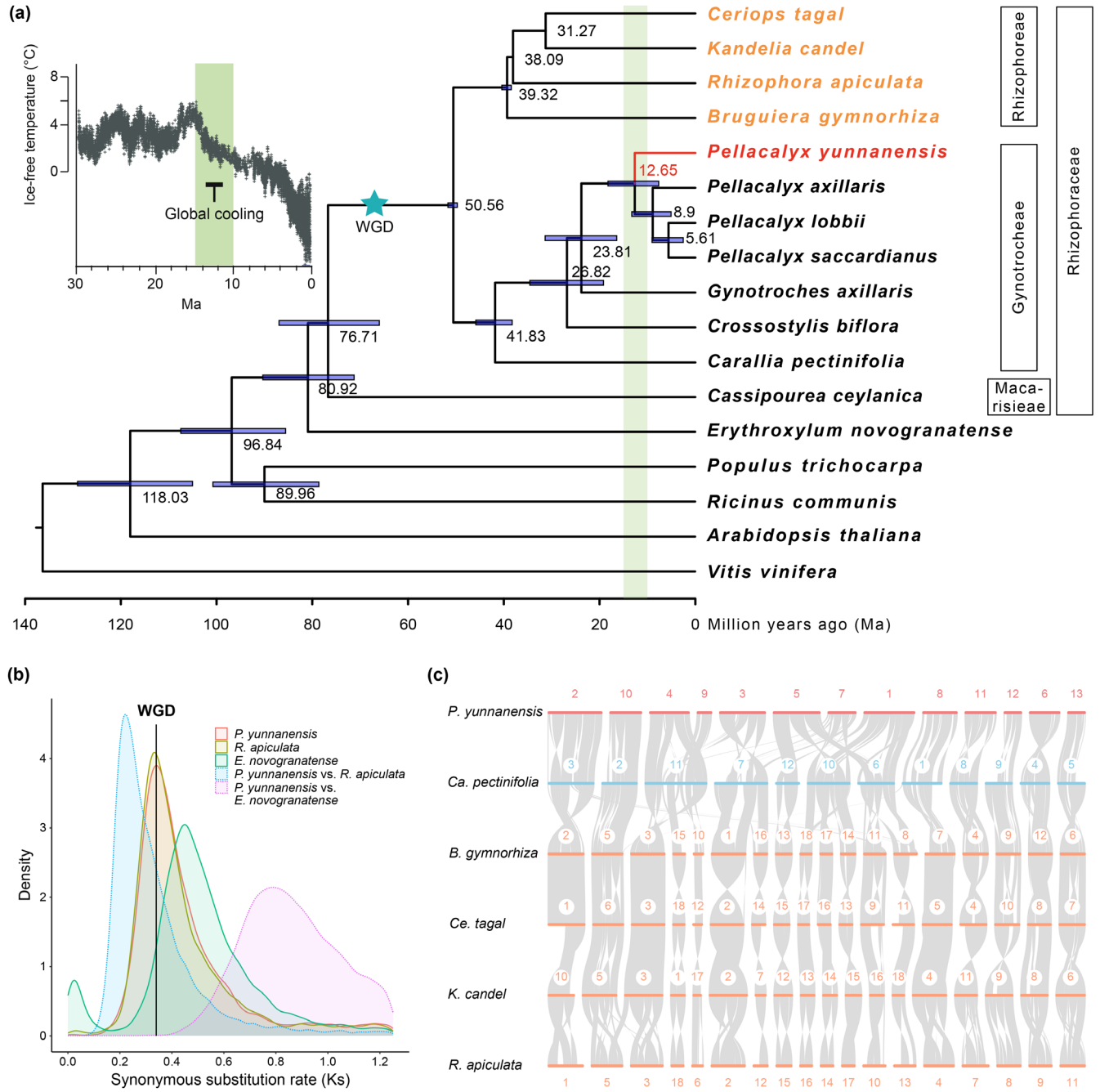

FIGURE 2 | Origin and chromosomal evolution of *P. yunnanensis*. (a) The origin of *P. yunnanensis* coincides with a global cooling period. The upper left panel shows historical temperature changes over 30 million years and highlights the global cooling period (modified from Zachos, Dickens, and Zeebe 2008). Divergence times were estimated using MCMCtree and the 95% confidence intervals were marked with blue bars. The whole-genome duplication (WGD) event was dated and marked with a star. (b) Synonymous substitution rate (Ks) distributions of intraspecific paralogs and interspecific orthologues. The Ks peak (0.34) of *P. yunnanensis* suggests a WGD event. (c) Orthologous collinear blocks in *P. yunnanensis*, *Ca. pectinifolia* and Rhizophoreae mangrove genomes. These collinear blocks were visualised by the 'jcvi.graphics.karyotype' program (Tang et al. 2008). Chromosome serial numbers are marked.

3.3 | Low Genetic Diversity and Adaptive Potential of *P. yunnanensis*

We aligned the DNA short reads to *P. yunnanensis* genome and identified 344,533 single nucleotide polymorphisms (SNPs) (Tables S7 and S8). The genome-wide heterozygosity is 1.20 per kb, much lower than nine inland plants and comparable to other endangered plants and domesticated crops (Figure 3a). The low genetic diversity suggests the low adaptive potential of *P. yunnanensis* in dealing with environmental changes.

Next, we used the pairwise sequentially Markovian coalescent (PSMC) model to infer the history of the *P. yunnanensis* population size. The model shows that the effective population size (N_e) of *P. yunnanensis* was similar to that of *Ca. pectinifolia* a million years ago and has decreased since then (Figure 3b). During the

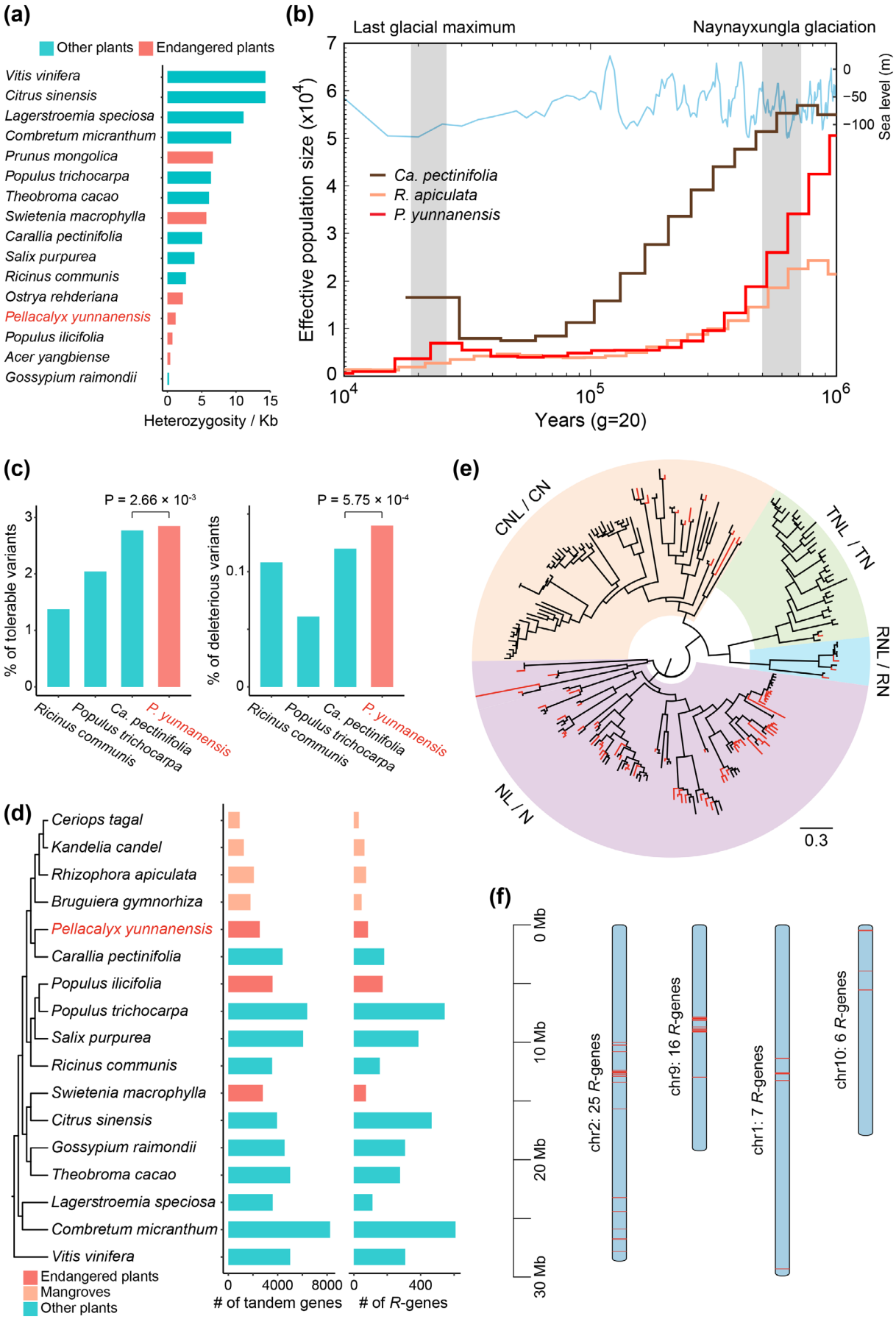

FIGURE 3 | Legend on next page.

FIGURE 3 | Low genetic diversity and adaptive potential of *P. yunnanensis*. (a) The heterozygosity of *P. yunnanensis* is lower than that of many inland plants and comparable to or even lower than sampled endangered species and domesticated crops. (b) Demographic histories of *P. yunnanensis*, *R. apiculata* and *Ca. pectinifolia*. Generation time (*n*) is set as 20 years for all species. Mutation rates of *Ca. pectinifolia*, *R. apiculata* and *P. yunnanensis* are set as 1.96×10^{-9}, 2.34×10^{-9} and 1.58×10^{-9} per site per year. Grey vertical bars represent the last glacial maximum (LGM) and Naynayxungla glaciation from left to right. (c) High proportion of deleterious mutations in *P. yunnanensis*. Fisher's exact test was used to determine significance. (d) Reduction of genes related to environmental adaptation in *P. yunnanensis* genome. Endangered species are marked in red, mangroves are marked in orange and other inland plants are marked in cyan. (e) Phylogeny of *P. yunnanensis* (red) and *Ca. pectinifolia R*-genes. (f) Distribution of *R*-genes in *P. yunnanensis* genome. Chromosomes with more than five *R*-genes (red lines) are shown.

largest Naynayxungla glaciation period on the Qinghai-Tibet Plateau (720 ka – 500 ka), N_e of *P. yunnanensis* decreased sharply, approaching the level of constantly 'endangered' mangroves (He et al. 2022; Zheng, Xu, and Shen 2002). Then, the *P. yunnanensis* N_e expanded slightly between 40 and 30 ka. By the last glacial maximum (LGM; 26.5 ka – 19 ka), the *P. yunnanensis* and Rhizophoreae mangrove population sizes experienced a second sharp decline (Clark et al. 2009). After LGM, the *P. yunnanensis* N_e dropped very low and has not recovered. These results suggest that *P. yunnanensis* population size is highly susceptible to changes in global climate.

Small population size usually results in low efficiency of natural selection and accelerated accumulation of deleterious mutations (Z. Chen et al. 2020). Non-synonymous substitutions (A) that cause changes in amino acid sequence are mainly deleterious, while synonymous substitutions (S) that do not affect protein composition are considered nearly neutral. Consistent with the prediction that small populations cannot effectively eliminate deleterious mutations, the A/S ratio is significantly higher in *P. yunnanensis* than in the *Ca. pectinifolia* relative (1.36 vs. 1.13; Fisher's exact test; $p < 10^{-15}$). To analyse patterns of genetic load, we classified deleterious variants based on their predicted impact on protein function as 'tolerable', such as likely to alter protein effectiveness, or 'deleterious', such as protein inactivation (Materials and Methods). Of the 379,529 amino acid changing variants in *P. yunnanensis*, 10,804 variants are classified as 'tolerable,' a significant increase compared with *Ca. pectinifolia* (46,098/1,663,649; Fisher's exact test; $p = 2.66\times10^{-3}$; Figure 3c; Table S8). The fraction of likely 'deleterious' substitutions is even higher than the control species, a 16.7% increase (0.14% vs. 0.12%; Fisher's exact test; $p = 5.75\times10^{-4}$). The pattern still holds when other Malpighiales plants are included for comparison. We identified 386 loss-of-function (LOF) mutations distributed in 339 genes, involving essential processes such as negative regulation of DNA and RNA metabolism (Figure S8). These results show that the *P. yunnanensis* genetic load is much higher than in its relatives, a highly unfavourable trend for its future survival.

In addition to genetic diversity, gene duplications also contribute to adaptive potential to environmental change (Van de Peer, Mizrachi, and Marchal 2017; Wu, Han, and Jiao 2019). *Pellacalyx yunnanensis* has 2400 fewer WGD-duplicated genes than *Ca. pectinifolia* (10,912 vs. 13,357). The patterns are even more striking for tandem duplications that deeply participate in environmental stress response. The number of tandem duplications of *P. yunnanensis* is half as prevalent as in the 10 inland relatives (2547 vs. 5078), sharing the same reduced trend as mangroves and other endangered species (Figure 3d). Thus, it appears that this avenue for meeting environmental challenges is also hampered in these species as incidence of duplicated genes can correlate with species vulnerability (Hamabata et al. 2019).

Among the tandem gene duplications, disease resistance genes (*R*-genes) play key roles in pathogeny resistance and are especially important for new environmental adaptation. The largest *R*-gene family contains the nucleotide-binding site (NBS) domain with the leucine-rich-repeat (LRR) domain often appearing at the C terminus. This family is further divided into four subclasses (NL/N, RNL/RN, TNL/TN and CNL/CN) according to whether the N terminus contains a Toll/Interleukin-1 receptor/resistance (TIR), coiled-coil (CC) or resistance to powdery mildew 8 (RPW8) domain (Liu et al. 2021). We found only 47% as many *R*-genes in *P. yunnanensis* as in *Ca. pectinifolia* (85 vs. 182), of which 77 contain only the NBS or NBS+LRR domain (Figure 3d; Table S9). The *R*-gene phylogeny of *P. yunnanensis* shows that three whole subclasses (RNL/RN, TNL/TN and CNL/CN) are nearly absent in this species (Figure 3e; Figure S9). In the largest NL/N subclass, 30% of *P. yunnanensis R*-genes arose due to recent tandem duplications (Figure 3f). Few *R*-genes, especially the significant reduction of some subclasses, may limit *P. yunnanensis*' adaptation to biotic stresses in changing environments.

3.4 | Molecular Mechanisms of Seed Non-dormancy

3.4.1 | Gene Family Evolution Contributes to the Non-dormant Seed of *P. yunnanensis*

Difficulty in population regeneration caused by seed non-dormancy may be an essential reason why *P. yunnanensis* is endangered (Ma et al. 1988). Considering that the seeds of other *Pellacalyx* species and the inland relative *Ca. pectinifolia* are physiologically dormant, *P. yunnanensis* ancestors may have lost this trait (Baskin and Baskin 2014). We examined the copy number evolution of 713 gene families related to seed development. We classify a gene family as reduced in *P. yunnanensis* if its size is less than half of other inland plants. Conversely, families that are twice as numerous as in other taxa are considered expanded. Using these criteria, we identified 15 lost, 41 reduced and 18 expanded gene families in *P. yunnanensis* (Figure 4a; Table S10).

Genes participating in the late stages of seed maturation experienced extensive copy number reduction in *P. yunnanensis*, in line with the loss of seed dormancy. ABA (Abscisic acid)-regulated late embryogenesis abundant (LEA) proteins that play a protective role in seed dehydration, such as *GEA6* and *SMP1*, are entirely lost (Figure 4c; Table S10) (Carles et al. 2002; Kushwaha et al. 2012). The *MPC*, *CP1* and *UMAMIT14* genes that take part

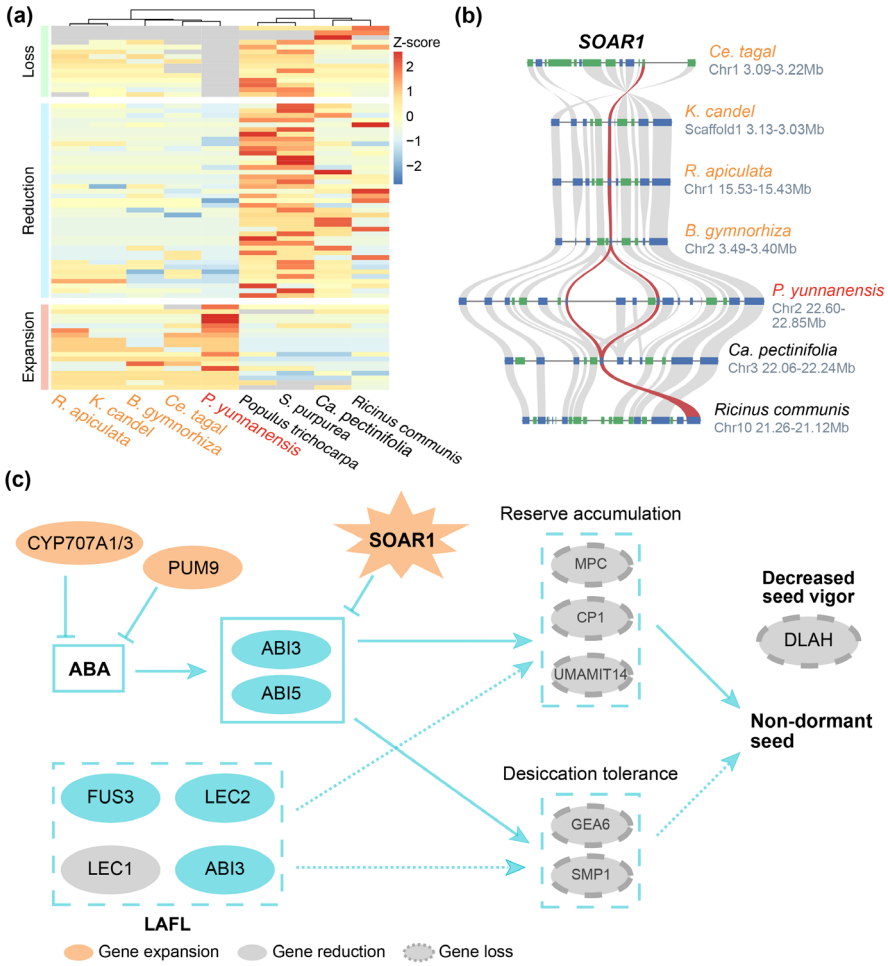

FIGURE 4 | Genetic changes associated with the seed dormancy loss in *P. yunnanensis*. (a) *P. yunnanensis* and Rhizophoreae mangroves exhibit correlated changes in the family size of seed development genes. Gene numbers are normalised using the Z-score; missing genes are represented in grey. The ward method of the R package "pheatmap" was used for clustering. (b) Interspecific orthologous collinear blocks of the *SOAR1* gene. The red lines show that the *SOAR1* gene is tandemly duplicated in *P. yunnanensis*. (c) Possible molecular mechanism of seed non-dormancy in *P. yunnanensis*. Arrows represent facilitation, bars represent inhibition and possible functional impairment is represented by dashed boxes and dashed arrows.

in the storage stage are also lost, resulting in seed reserve accumulation defects and tiny seeds observed in *P. yunnanensis* (Müller et al. 2015; Qiu, Liu, and Adams 2014; Tsuji et al. 2013). The *DLAH* gene mutation can significantly reduce seed vigour, and its loss may be responsible for the rapid loss of *P. yunnanensis* seed germination success (Seo, Kim, and Kim 2011).

The ABA signalling pathway is a central player in promoting seed dormancy. We found significantly altered gene families that are deeply involved in these processes. The *SOAR1* gene is a key factor that negatively regulates ABA signalling pathway. It expanded tandemly to two copies in *P. yunnanensis* but is present in a single copy in 12 inland relatives (Figure 4b; Table S11). RNA-seq data confirmed that both copies are expressed (Figure S10). As an upstream regulator of ABI5, overexpression of *SOAR1* renders seeds and seedlings unresponsive to ABA, promoting early germination (Figure 4c) (Mei et al. 2014). The expansion of gene families involved in ABA degradation, such as *CYP707A1* and *PUM9*, may also inhibit ABA signalling (Nyikó, Auber, and Bucher 2019; Okamoto et al. 2006). LAFL (FUS3, ABI3, LEC1/L1L and LEC2) is another key seed development regulatory network independent of the ABA pathway. Among them, the *LEC1* gene has only one copy in *P. yunnanensis*, while its relatives in Malpighiales all have two copies (Table S10). Mutations in all or part of the LAFL network may cause seeds to skip dormancy and germinate earlier (K. Chen et al. 2020). The inhibition of the ABA pathway mainly caused by the expansion of the *SOAR1* family and the impaired function of the LAFL network may have contributed to the loss of seed dormancy in *P. yunnanensis*.

3.4.2 | Independent Loss of Seed Dormancy in *P. yunnanensis* and Rhizophoreae Mangroves

Rhizophoreae mangroves, relatives of *P. yunnanensis*, have also lost seed dormancy (Tomlinson 2016). These mangrove seeds, known as viviparous seeds, germinate even earlier while still attached to the mother plant. Considering that *P. yunnanensis* and Rhizophoreae mangroves are two independent clades with close kinship, it is worth exploring whether these taxa lost seed dormancy through a convergent mechanism.

There are many convergent changes in *P. yunnanensis* and Rhizophoreae mangroves, with 34 gene families involved in

seed germination showing correlated expansion or contraction (Figure 4a; Table S12). *GEA6* and *SMP1* genes were convergently lost in *P. yunnanensis* and Rhizophoreae mangroves. They are LEA proteins and should act as protective substances when seeds are dehydrated. The *CYP707A1* gene involved in ABA catabolism retained both duplicated copies after WGD. The other 31, genes are mainly involved in organic matter synthesis, ion transport, protein localization and cell wall production.

The significant difference between the two clades is the upstream regulation of ABA signalling (Figure S11). As shown above, the repression of ABA signalling in *P. yunnanensis* may have been achieved by *SOAR1* gene duplication. However, *SOAR1* did not expand in the ancestors of the Rhizophoreae mangroves, and only two of the eight surveyed species have recent duplications (Table S11). Previous studies suggested the loss of *DOG1* gene, which promotes the ABA signalling pathway, has contributed to the vivipary of Rhizophoreae mangroves. Deleting *DOG1* downregulates upstream ABA-responsive genes such as *ABI3* and *ABI5*, causing seeds to lose dormancy (K. Chen et al. 2020; Xu et al. 2024).

3.5 | Genomic Divergence Between Inland Plants and Rhizophoreae Mangroves

The *P. yunnanensis* genome also offers opportunities to investigate the genomic changes associated with adaptation to terrestrial or intertidal environments. The ancestors of Rhizophoreae mangroves and Gynotrocheae (*P. yunnanensis* and *Ca. pectinifolia*) experienced a WGD event during the K-Pg boundary period (Figure 2a). We test whether differences in genes retained after the WGD may reflect adaptation to distinct environments.

We identified 3327–5330 WGD duplication gene pairs across the six species (Figure 5a; Figure S12). On average, 61.9% of the WGD duplications are conserved in all species. We consider gene families present in more than half of the four mangrove species but absent in all inland plants as unique among mangroves; otherwise, they are defined as unique in inland plants. As a result, there are 342 WGD duplications unique among mangroves and 781 such loci in inland plants.

Conserved WGD duplications among all species are mainly involved in organic substance metabolic processes, organic substance biosynthetic processes, developmental regulation and stress response. Consistent with the trends shown in many reports, these loci may help these plants face multiple environmental perturbations during the K-Pg boundary period (Figure 5b) (Feng et al. 2024; Van de Peer, Mizrachi, and Marchal 2017; Xu et al. 2017). The WGD duplications unique in Rhizophoreae mangroves are involved in post-embryonic and flower development as well as cell growth (Table S13). Specifically, *IAA28/30* and *LBD29* promote aerial root growth; *SOS5* and *AHA10* genes are involved in seed coat growth; and *GA20OX4* responds to gibberellins to promote germination (Appelhagen et al. 2015; Basu et al. 2016; Rieu et al. 2008; Zhang et al. 2020). The WGD duplications unique in Gynotrocheae inland plants involve macromolecule biosynthetic processes, intracellular transport and cell localization establishment (Table S14). In particular, the salicylic acid mediated signalling pathway expansion may increase their resistance to biotic stress (Huot et al. 2014). The differences in retained genes reflect distinct adaptations to terrestrial and intertidal environments.

Aerial roots are a common trait among Rhizophoreae mangroves and Gynotrocheae inland plants. It further evolved in Rhizophoreae mangroves to resist frequent flooding in intertidal zones. We found 14 WGD-duplicated gene families involved in aerial root development in both tribes (Figure 5c; Table S15) (Zhang et al. 2020). Under light stimulation, the photoreceptor *PHR2* drives *TAR2* to produce high levels of auxin in roots. High auxin levels then promote aerial root initiation through the TCP–SCR–PLT complex and drive cell cycle activation and cell patterning through *LBD29*. In addition, PLT and SHR proteins establish quiescent centres and play an essential role in the aerial root apical meristem. We also found three unique WGD duplications in Rhizophoreae mangroves. Among them, the *ARF19* transcription factor can enhance the expression of *LBD29*, while the *IAA28* and *IAA30* promote lateral root generation and regulate quiescent centres, respectively (Figure 5c,d). Aerial roots in Rhizophoreae mangroves and Gynotrocheae are formed through a light-stimulated auxin-dependent pathway, and mangroves specialise in this trait through unique WGD duplications.

4 | Discussion

Understanding the molecular mechanisms underlying endangerment is crucial for conserving endangered plants, especially PSESPs (Ma et al. 2013; Pimm et al. 2014; Sun, Yang, and Dao 2019; Yang et al. 2020). In this study, we take the case of the *Pellacalyx yunnanensis*, a critically endangered plant with extremely small populations exclusively found in Xishuangbanna, China, to demonstrate how a chromosome-level genome can facilitate conservation efforts. We inferred *P. yunnanensis*' origin and evolutionary history, estimated its population size changes and evaluated this species' adaptive potential when faced with future climate change, all providing clues for conserving PSESPs.

The habitat of *P. yunnanensis* is the tropical rainforest at an altitude of 650–800 m in southern Yunnan, unlike other Rhizophoraceae found in coastal areas (Juncosa and Tomlinson 1988a; Wu and Raven 1994). Using genomic data, we found that the species has experienced several rounds of population reduction resulting from historical climate changes. According to geological data, southern Yunnan was on the east coast of the Tethys Sea in the Palaeocene, which was suitable for reproducing Rhizophoraceae species (Yunnan Bureau of Geology and Mineral Resources 1995). In the Eocene, starting with the Himalayan orogeny, the crust of southern Yunnan was uplifted and gradually formed the current mountainous terrain. Subsidence of the south Yunnan crust during the global cooling in the Miocene formed many lake basins with a warm and humid climate (Zachos et al. 2001; Zhu 2022). *Pellacalyx yunnanensis* may have remained in this warm refugium. The high altitude of the area around Xishuangbanna limited its spread and resulted in a more severe population size reduction during historical climate changes (Figure 1a). The further reduction in *P. yunnanensis* N_e coincided with the emergence of Naynayxungla,

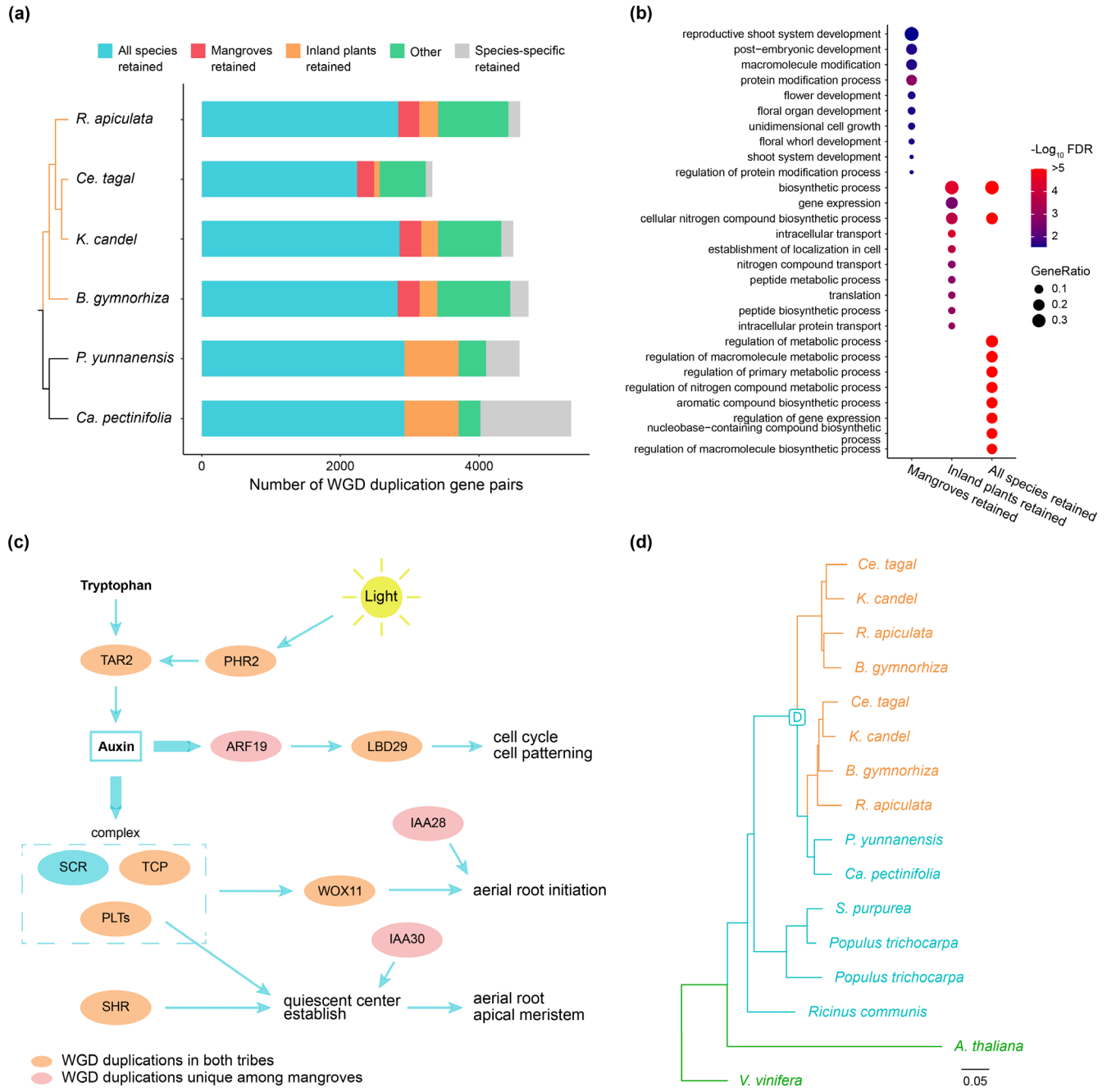

FIGURE 5 | Whole genome duplication evolution (WGD) and its contribution to mangrove aerial root formation. (a) WGD duplications in Rhizophoreae mangroves (orange) and inland species. (b) GO enrichment analyses of WGD duplications. The top 10 enriched GO terms of each WGD duplication category are shown. (c) WGD duplications are involved in auxin accumulation and aerial root development. (d) Phylogenetic tree of the *ARF19* gene family. A box marks the node where the WGD event occurred. Mangroves, inland plants of Malpighiales and outgroups are marked in orange, blue and green, respectively.

the largest glacial period in the Qinghai-Tibet Plateau during the Quaternary (Zheng, Xu, and Shen 2002). It fell extremely low after the LGM (Clark et al. 2009). The ancestor plants with wide distribution have populations trapped in refugium during historical climate changes, which would be a universal mechanism for forming PSESPs (Ma et al. 2013; Yang et al. 2020). The population sizes were constrained by geographical conditions and more easily impacted by climate changes.

Difficulty in reproduction is one of the major reasons causing endangerment of *P. yunnanensis* and other PSESPs. We found incompatibility between seed germination and the local environments. The species produce many seeds but the survival rate is extremely low, coupled with harsh seedling growth conditions (Ma et al. 1988). Loss of seed dormancy may have been advantageous for *P. yunnanensis* ancestors but has become an important reason for its endangered status. We found that the independent loss of seed dormancy in *P. yunnanensis* and Rhizophoreae mangroves may result from gene family evolution. Both clades have lost many genes involved in seed maturation and material storage, resulting in smaller seeds and a rapid decline in vigour. The major difference between the two clades occurred in the

repression of ABA signalling that promotes seed dormancy. Rhizophoreae mangroves have lost the *DOG1* gene, a positive regulator of ABA signalling (Xu et al. 2024). *Pellacalyx yunnanensis* has retained the *DOG1* gene but expanded the *SOAR1* gene that negatively regulates ABA signalling.

We found that *P. yunnanensis* has struggled to survive in the past and may also have a challenging future. Genomic features suggest that this species may have a low adaptive potential in the face of future climate changes or new environments. The genetic diversity is low, while deleterious mutations have accumulated, possibly leading to inbreeding depression. We also found fewer duplicated genes, especially those involved in disease resistance, making it difficult to cope with various biological stresses. Finally, many genes promoting seed dormancy are also involved in drought resistance (Jiang et al. 2015; Li et al. 2017). The loss of such genes may result in an inability to adapt to drought.

Although *P. yunnanensis* has experienced dramatic population decline and lives in unfavoured climates, it could still survive for millions of years. However, human activities such as planting economic forests have nearly made the species extinct in less than 100 years (Su et al. 2005). Protection efforts thus must pay special attention to increasing genetic diversity while limiting adverse human intervention. Based on the study's results, we suggested in situ conservation should be a priority instead of ex situ conservation, because *P. yunnanensis* has low adaptability to new environments. We found that seedling breeding should be an efficient way to improve reproduction because the environments did not meet seed germination requirements. The molecular mechanism analysis of the non-dormant seeds provides us with the prospect of using technologies like gene editing to improve *P. yunnanensis* seeds in the future, making them adaptable to the local climate.

The assembled chromosome-level genome of *P. yunnanensis* also helps study the origin and genome evolution of the Rhizophoraceae (Schwarzbach and Ricklefs 2000). For example, according to the extensive intra-chromosomal collinear blocks in *P. yunnanensis* and *Ca. pectinifolia*, we could infer that the chromosome number difference between the Gynotrocheae and the Rhizophoreae mangroves should be the result of chromosome fusion in the Gynotrocheae (Oginuma and Tobe 1997; Senakun 2018). In the future, supplementing high-throughput chromosome conformation capture (Hi-C) data would further improve assembly quality and promote the identification of more complete and more accurate genome structural variations that might be related to speciation and adaptive evolution.

In summary, we explored the evolutionary history of *P. yunnanensis* and determined that its origin and population decline are primarily the result of an inability to adapt to a cooling climate. We found that this species has low genetic diversity, genome-wide accumulation of deleterious mutations and extensively reduced tandem gene duplications and disease resistance genes. All of these genomic features suggest the low adaptation potential to future climate change. Finally, the expansion of the *SOAR1* gene that inhibits the ABA signal, the streamlining of the LAFL network and the extensive loss of genes related to reserve accumulation and desiccation tolerance may be the key to seed non-dormancy. These findings would guide policy formulation of this extremely endangered plant. Given its low adaptation potential, we suggest that in situ conservation should be a priority for *P. yunnanensis*. Seedling breeding should be an efficient way to improve reproduction. The revelation of molecular mechanisms underlying non-dormant seeds allows for methods such as gene editing to improve seeds and alleviate the reproduction difficulties of *P. yunnanensis*. Our study provides an example of applying genomic data to PSESP conservation.

Author Contributions

S. Shi and S.X. conceived and designed the research. S. Shao, S.X., Y.L., X.F., C.J., M.L., R.Z., M.E.T. and Z.G. collected materials. S. Shao and S.X. performed the experiments. S. Shao and S.X. carried out data processing and analysis. S. Shao, S.X., Z.H. and S. Shi wrote the first draft of the manuscript. All authors read and approved the final version.

Acknowledgements

We thank Chung-I Wu from Sun Yat-sen University for the helpful discussion. This study was supported by the National Natural Science Foundation of China (32330005, 32100176 and 32370227), the Guangdong Basic and Applied Basic Research Foundation (2022A1515010962 and 2023B1515020083) and the China Postdoctoral Science Foundation (2024T170174).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The genome assembly sequences and raw Illumina reads have been deposited in the National Center for Biotechnology Information (NCBI) with BioProject number PRJNA1043905. The genome annotations are available at Figshare (https://doi.org/10.6084/m9.figshare.27159180).

References

Appelhagen, I., N. Nordholt, T. Seidel, et al. 2015. "TRANSPARENT TESTA 13 Is a Tonoplast P_{3A}-ATPase Required for Vacuolar Deposition of Proanthocyanidins in *Arabidopsis thaliana* Seeds." *Plant Journal* 82, no. 5: 840–849.

Baskin, C. C., and J. M. Baskin. 2014. *Seeds: Ecology, Biogeography, and, Evolution of Dormancy and Germination*. San Diego: Academic Press.

Basu, D., L. Tian, T. DeBrosse, et al. 2016. "Glycosylation of a Fasciclin-Like Arabinogalactan-Protein (SOS5) Mediates Root Growth and Seed Mucilage Adherence via a Cell Wall Receptor-Like Kinase (FEI1/FEI2) Pathway in Arabidopsis." *PLoS One* 11, no. 1: 1–27.

Birney, E., M. Clamp, and R. Durbin. 2004. "GeneWise and Genomewise." *Genome Research* 14, no. 5: 988–995.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30, no. 15: 2114–2120.

Bredeson, J. V., J. B. Lyons, S. E. Prochnik, et al. 2016. "Sequencing Wild and Cultivated Cassava and Related Species Reveals Extensive Interspecific Hybridization and Genetic Diversity." *Nature Biotechnology* 34, no. 5: 562–570.

Buchfink, B., C. Xie, and D. H. Huson. 2014. "Fast and Sensitive Protein Alignment Using DIAMOND." *Nature Methods* 12, no. 1: 59–60.

Camacho, C., G. Coulouris, V. Avagyan, et al. 2009. "BLAST+: Architecture and Applications." *BMC Bioinformatics* 10, no. 1: 421.

Carles, C., N. Bies-Etheve, L. Aspart, et al. 2002. "Regulation of *Arabidopsis thaliana Em* Genes: Role of ABI5." *Plant Journal* 30, no. 3: 373–383.

Castresana, J. 2000. "Selection of Conserved Blocks From Multiple Alignments for Their Use in Phylogenetic Analysis." *Molecular Biology and Evolution* 17, no. 4: 540–552.

Chan, A. P., J. Crabtree, Q. Zhao, et al. 2010. "Draft Genome Sequence of the Oilseed Species *Ricinus communis*." *Nature Biotechnology* 28, no. 9: 951–956.

Chen, C., Y. Wu, J. Li, et al. 2023. "TBtools-II: A "One for all, all for One" Bioinformatics Platform for Biological Big-Data Mining." *Molecular Plant* 16, no. 11: 1733–1742.

Chen, K., G. J. Li, R. A. Bressan, C. P. Song, J. K. Zhu, and Y. Zhao. 2020. "Abscisic Acid Dynamics, Signaling, and Functions in Plants." *Journal of Integrative Plant Biology* 62, no. 1: 25–54.

Chen, Z., F. Ai, J. Zhang, et al. 2020. "Survival in the Tropics Despite Isolation, Inbreeding and Asexual Reproduction: Insights From the Genome of the World's Southernmost Poplar (*Populus ilicifolia*)." *Plant Journal* 103, no. 1: 430–442.

Cheng, H., G. T. Concepcion, X. Feng, H. Zhang, and H. Li. 2021. "Haplotype-Resolved De Novo Assembly Using Phased Assembly Graphs With Hifiasm." *Nature Methods* 18, no. 2: 170–175.

Cingolani, P., A. Platts, L. L. Wang, et al. 2012. "A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of *Drosophila melanogaster* Strain w^{1118}; *iso-2*; *iso-3*." *Fly* 6, no. 2: 80–92.

Clark, P. U., A. S. Dyke, J. D. Shakun, et al. 2009. "The Last Glacial Maximum." *Science* 325, no. 5941: 710–714.

Danielson, J. J., and D. B. Gesch. 2011. "Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010)." *U.S. Geological Survey*, no. 2011-1073.

Depristo, M. A., E. Banks, R. Poplin, et al. 2011. "A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data." *Nature Genetics* 43, no. 5: 491–501.

Díaz, S., J. Settele, E. S. Brondízio, et al. 2019. "Pervasive Human-Driven Decline of Life on Earth Points to the Need for Transformative Change." *Science* 366, no. 6471: eaax3100.

Doyle, J. J., and J. L. Doyle. 1987. "A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue." *Phytochemical Bulletin* 19: 11–15.

Duke, N. C. 1992. "Mangrove Floristics and Biogeography." In *Tropical Mangrove Ecosystems*, 63–100. Washington, DC: American Geophysical Union.

Emms, D. M., and S. Kelly. 2019. "OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics." *Genome Biology* 20, no. 1: 1–14.

Feng, X., Q. Chen, W. Wu, et al. 2024. "Genomic Evidence for Rediploidization and Adaptive Evolution Following the Whole-Genome Triplication." *Nature Communications* 15, no. 1: 1635.

Flynn, J. M., R. Hubley, C. Goubert, et al. 2020. "RepeatModeler2 for Automated Genomic Discovery of Transposable Element Families." *Proceedings of the National Academy of Sciences of the United States of America* 117, no. 17: 9451–9457.

Graham, A. 2006. "Paleobotanical Evidence and Molecular Data in Reconstructing the Historical Phytogeography of Rhizophoraceae." *Annals of the Missouri Botanical Garden* 93, no. 2: 325–334.

Haas, B. J., S. L. Salzberg, W. Zhu, et al. 2008. "Automated Eukaryotic Gene Structure Annotation Using EVidenceModeler and the Program to Assemble Spliced Alignments." *Genome Biology* 9, no. 1: 1–22.

Hamabata, T., G. Kinoshita, K. Kurita, et al. 2019. "Endangered Island Endemic Plants Have Vulnerable Genomes." *Communications Biology* 2, no. 1: 1–10.

He, Z., X. Feng, Q. Chen, et al. 2022. "Evolution of Coastal Forests Based on a Full Set of Mangrove Genomes." *Nature Ecology and Evolution* 6, no. 6: 738–749.

Hellsten, U., K. M. Wright, J. Jenkins, et al. 2013. "Fine-Scale Variation in Meiotic Recombination in *Mimulus* Inferred From Population Shotgun Sequencing." *Proceedings of the National Academy of Sciences of the United States of America* 110, no. 48: 19478–19482.

Huot, B., J. Yao, B. L. Montgomery, and S. Y. He. 2014. "Growth-Defense Tradeoffs in Plants: A Balancing Act to Optimize Fitness." *Molecular Plant* 7, no. 8: 1267–1287.

Jiang, S. C., C. Mei, S. Liang, et al. 2015. "Crucial Roles of the Pentatricopeptide Repeat Protein SOAR1 in *Arabidopsis* Response to Drought, Salt and Cold Stresses." *Plant Molecular Biology* 88, no. 4–5: 369–385.

Jones, P., D. Binns, H. Y. Chang, et al. 2014. "InterProScan 5: Genome-Scale Protein Function Classification." *Bioinformatics* 30, no. 9: 1236–1240.

Juncosa, A. M., and P. B. Tomlinson. 1988a. "A Historical and Taxonomic Synopsis of Rhizophoraceae and Anisophylleaceae." *Annals of the Missouri Botanical Garden* 75, no. 4: 1278.

Juncosa, A. M., and P. B. Tomlinson. 1988b. "Systematic Comparison and Some Biological Characteristics of Rhizophoraceae and Anisophylleaceae." *Annals of the Missouri Botanical Garden* 75, no. 4: 1296–1318.

Katoh, K., and D. M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30, no. 4: 772–780.

Kent, W. J. 2002. "BLAT—The BLAST-Like Alignment Tool." *Genome Research* 12, no. 4: 656–664.

Kim, D., J. M. Paggi, C. Park, C. Bennett, and S. L. Salzberg. 2019. "Graph-Based Genome Alignment and Genotyping With HISAT2 and HISAT-Genotype." *Nature Biotechnology* 37, no. 8: 907–915.

Kovaka, S., A. V. Zimin, G. M. Pertea, R. Razaghi, S. L. Salzberg, and M. Pertea. 2019. "Transcriptome Assembly From Long-Read RNA-Seq Alignments With StringTie2." *Genome Biology* 20, no. 1: 1–13.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. "RAxML-NG: A Fast, Scalable and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference." *Bioinformatics* 35, no. 21: 4453–4455.

Krzywinski, M., J. Schein, I. Birol, et al. 2009. "Circos: An Information Aesthetic for Comparative Genomics." *Genome Research* 19, no. 9: 1639–1645.

Kushwaha, R., T. D. Lloyd, K. R. Schäfermeyer, S. Kumar, and A. B. Downie. 2012. "Identification of Late Embryogenesis Abundant (LEA) Protein Putative Interactors Using Phage Display." *International Journal of Molecular Sciences* 13, no. 6: 6582–6603.

Lamesch, P., T. Z. Berardini, D. Li, et al. 2012. "The Arabidopsis Information Resource (TAIR): Improved Gene Annotation and New Tools." *Nucleic Acids Research* 40, no. D1: D1202–D1210.

Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform." *Bioinformatics* 25, no. 14: 1754–1760.

Li, H., and R. Durbin. 2011. "Inference of Human Population History From Individual Whole-Genome Sequences." *Nature* 475, no. 7357: 493–496.

Li, H., B. Handsaker, A. Wysoker, et al. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25, no. 16: 2078–2079.

Li, X., Y. Zhang, L. Yin, and J. Lu. 2017. "Overexpression of Pathogen-Induced Grapevine TIR-NB-LRR Gene *VaRGA1* Enhances Disease Resistance and Drought and Salt Tolerance in *Nicotiana benthamiana*." *Protoplasma* 254, no. 2: 957–969.

Liu, Y., Z. Zeng, Y. M. Zhang, et al. 2021. "An Angiosperm NLR Atlas Reveals That NLR Gene Reduction Is Associated With Ecological Specialization and Signal Transduction Component Deletion." *Molecular Plant* 14, no. 12: 2015–2031.

Lomsadze, A., V. Ter-Hovhannisyan, Y. O. Chernoff, and M. Borodovsky. 2005. "Gene Identification in Novel Eukaryotic Genomes by Self-Training Algorithm." *Nucleic Acids Research* 33, no. 20: 6494–6506.

Ma, H., Y. Liu, D. Liu, et al. 2021. "Chromosome-Level Genome Assembly and Population Genetic Analysis of a Critically Endangered Rhododendron Provide Insights Into Its Conservation." *Plant Journal* 107, no. 5: 1533–1545.

Ma, X., Y. Zhuliang, X. Zaifu, and T. Guoda. 1988. "A Study of the Causes Threatening *Pellacalyx yunnanensis*, a Species Receiving Priority Protection in China." *Plant Diversity* 10, no. 3: 311–316.

Ma, Y., G. Chen, R. Edward Grumbine, Z. Dao, W. Sun, and H. Guo. 2013. "Conserving Plant Species With Extremely Small Populations (PSESP) in China." *Biodiversity and Conservation* 22, no. 3: 803–809.

Manni, M., M. R. Berkeley, M. Seppey, F. A. Simão, and E. M. Zdobnov. 2021. "BUSCO Update: Novel and Streamlined Workflows Along With Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes." *Molecular Biology and Evolution* 38, no. 10: 4647–4654.

McKenna, A., M. Hanna, E. Banks, et al. 2010. "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data." *Genome Research* 20, no. 9: 1297–1303.

Mei, C., S. C. Jiang, Y. F. Lu, et al. 2014. "*Arabidopsis* Pentatricopeptide Repeat Protein SOAR1 Plays a Critical Role in Abscisic Acid Signalling." *Journal of Experimental Botany* 65, no. 18: 5317–5330.

Mistry, J., S. Chuguransky, L. Williams, et al. 2021. "Pfam: The Protein Families Database in 2021." *Nucleic Acids Research* 49, no. D1: D412–D419.

Morris, J. L., M. N. Puttick, J. W. Clark, et al. 2018. "The Timescale of Early Land Plant Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 10: E2274–E2283.

Motamayor, J. C., K. Mockaitis, J. Schmutz, et al. 2013. "The Genome Sequence of the Most Widely Cultivated Cacao Type and Its Use to Identify Candidate Genes Regulating Pod Color." *Genome Biology* 14, no. 6: 1–25.

Müller, B., A. Fastner, J. Karmann, et al. 2015. "Amino Acid Export in Developing *Arabidopsis* Seeds Depends on UmamiT Facilitators." *Current Biology* 25, no. 23: 3126–3131.

Muller, J. 1981. "Fossil Pollen Records of Extant Angiosperms." *Botanical Review* 47, no. 1: 1–142.

Nyikó, T., A. Auber, and E. Bucher. 2019. "Functional and Molecular Characterization of the Conserved Arabidopsis PUMILIO Protein, APUM9." *Plant Molecular Biology* 100, no. 1: 199–214.

Oginuma, K., and H. Tobe. 1997. "Chromosomes of *Carallia*, *Gynotroches*, and *Pellacalyx* (Rhizophoraceae; Gynotrocheae)." *Acta Phytotaxonomica et Geobotanica* 48, no. 1: 15–21.

Okamoto, M., A. Kuwahara, M. Seo, et al. 2006. "CYP707A1 and CYP707A2, Which Encode Abscisic Acid 8′-Hydroxylases, Are Indispensable for Proper Control of Seed Dormancy and Germination in Arabidopsis." *Plant Physiology* 141, no. 1: 97–107.

Ou, S., and N. Jiang. 2018. "LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons." *Plant Physiology* 176, no. 2: 1410–1422.

Ouyang, S., W. Zhu, J. Hamilton, et al. 2007. "The TIGR Rice Genome Annotation Resource: Improvements and New Features." *Nucleic Acids Research* 35, no. Suppl 1: D883–D887.

Paez, S., R. H. S. Kraus, B. Shapiro, M. T. P. Gilbert, and E. D. Jarvis. 2022. "Reference Genomes for Conservation." *Science* 377, no. 6604: 364–366.

Pimm, S. L., C. N. Jenkins, R. Abell, et al. 2014. "The Biodiversity of Species and Their Rates of Extinction, Distribution, and Protection." *Science* 344, no. 6187: 1246752.

Qin, H., Y. Yang, S. Dong, et al. 2017. "Threatened Species List of China's Higher Plants." *Biodiversity Science* 25, no. 7: 696–744.

Qiu, Y., S. L. Liu, and K. L. Adams. 2014. "Frequent Changes in Expression Profile and Accelerated Sequence Evolution of Duplicated Imprinted Genes in *Arabidopsis*." *Genome Biology and Evolution* 6, no. 7: 1830–1842.

Renne, P. R., A. L. Deino, F. J. Hilgen, et al. 2013. "Time Scales of Critical Events Around the Cretaceous-Paleogene Boundary." *Science* 339, no. 6120: 684–687.

Rieu, I., O. Ruiz-Rivero, N. Fernandez-Garcia, et al. 2008. "The Gibberellin Biosynthetic Genes *AtGA20ox1* and *AtGA20ox2* Act, Partially Redundantly, to Promote Growth and Development Throughout the Arabidopsis Life Cycle." *Plant Journal* 53, no. 3: 488–504.

Schwarzbach, A. E., and R. E. Ricklefs. 2000. "Systematic Affinities of Rhizophoraceae and Anisophylleaceae, and Intergeneric Relationships Within Rhizophoraceae, Based on Chloroplast DNA, Nuclear Ribosomal DNA, and Morphology." *American Journal of Botany* 87, no. 4: 547–564.

Senakun, C. 2018. "Chromosome Numbers for Selected Thailand Plant Species." *Thai Forest Bulletin (Botany)* 46, no. 1: 67–71.

Seo, Y. S., E. Y. Kim, and W. T. Kim. 2011. "The *Arabidopsis sn*-1-Specific Mitochondrial Acylhydrolase AtDLAH Is Positively Correlated With Seed Viability." *Journal of Experimental Botany* 62, no. 15: 5683–5698.

Seppey, M., M. Manni, and E. M. Zdobnov. 2019. "BUSCO: Assessing Genome Assembly and Annotation Completeness." *Methods in Molecular Biology* 1962: 227–245.

Setoguchi, H., K. Kosuge, and H. Tobe. 1999. "Molecular Phylogeny of Rhizophoraceae Based on *rbcL* Gene Sequences." *Journal of Plant Research* 112, no. 4: 443–455.

Shi, S., Y. Huang, K. Zeng, et al. 2005. "Molecular Phylogenetic Analysis of Mangroves: Independent Evolutionary Origins of Vivipary and Salt Secretion." *Molecular Phylogenetics and Evolution* 34, no. 1: 159–166.

Shi, S., Y. Zhong, Y. Huang, Y. Du, X. Qiu, and H. Chang. 2002. "Phylogenetic Relationships of the Rhizophoraceae in China Based on Sequences of the Chloroplast Gene *matK* and the Internal Transcribed Spacer Regions of Nuclear Ribosomal DNA and Combined Data Set." *Biochemical Systematics and Ecology* 30, no. 4: 309–319.

Slater, G. S. C., and E. Birney. 2005. "Automated Generation of Heuristics for Biological Sequence Comparison." *BMC Bioinformatics* 6, no. 1: 31.

Stanke, M., O. Keller, I. Gunduz, A. Hayes, S. Waack, and B. Morgenstern. 2006. "AUGUSTUS: *Ab Initio* Prediction of Alternative Transcripts." *Nucleic Acids Research* 34, no. Web Server issue: W435–W439.

Su, Z., S. Yin, C. Wu, and Z. Cheng. 2005. "Analysis of Genetic Structure of the Endangered Species *Pellacalyx yunnanensis* (Rhizophoraceae) by RAPD." *Acta Botanica Yunnanica* 27, no. 2: 181–186.

Sun, W., J. Yang, and Z. Dao. 2019. *Study and Conservation of Plant Species With Extremely Small Populations (PSESP) in Yunnan Province, China*. Beijing: Science Press.

Suyama, M., D. Torrents, and P. Bork. 2006. "PAL2NAL: Robust Conversion of Protein Sequence Alignments Into the Corresponding Codon Alignments." *Nucleic Acids Research* 34, no. Web Server issue: W609–W612.

Tang, H., J. E. Bowers, X. Wang, R. Ming, M. Alam, and A. H. Paterson. 2008. "Synteny and Collinearity in Plant Genomes." *Science* 320, no. 5875: 486–488.

Tang, H., V. Krishnakumar, S. Bidwell, et al. 2014. "An Improved Genome Release (Version Mt4.0) for the Model Legume *Medicago truncatula*." *BMC Genomics* 15, no. 1: 1–14.

Tarailo-Graovac, M., and N. Chen. 2009. "Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences." *Current Protocols in Bioinformatics* 25, no. Suppl 25: 4.10.1–4.10.14.

The French–Italian Public Consortium for Grapevine Genome Characterization. 2007. "The Grapevine Genome Sequence Suggests Ancestral Hexaploidization in Major Angiosperm Phyla." *Nature* 449, no. 7161: 463–467.

Tomlinson, P. B. 2016. *The Botany of Mangroves*. Cambridge: Cambridge University Press.

Tsuji, A., K. Tsukamoto, K. Iwamoto, Y. Ito, and K. Yuasa. 2013. "Enzymatic Characterization of Germination-Specific Cysteine Protease-1 Expressed Transiently in Cotyledons During the Early Phase of Germination." *Journal of Biochemistry* 153, no. 1: 73–83.

Tuskan, G. A., S. DiFazio, S. Jansson, et al. 2006. "The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray)." *Science* 313, no. 5793: 1596–1604.

Urban, M. C. 2015. "Accelerating Extinction Risk From Climate Change." *Science* 348, no. 6234: 571–573.

Van de Peer, Y., E. Mizrachi, and K. Marchal. 2017. "The Evolutionary Significance of Polyploidy." *Nature Reviews Genetics* 18, no. 7: 411–424.

Vurture, G. W., F. J. Sedlazeck, M. Nattestad, et al. 2017. "GenomeScope: Fast Reference-Free Genome Profiling From Short Reads." *Bioinformatics* 33, no. 14: 2202–2204.

Walker, B. J., T. Abeel, T. Shea, et al. 2014. "Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement." *PLoS One* 9, no. 11: e112963.

Wang, D., Y. Zhang, Z. Zhang, J. Zhu, and J. Yu. 2010. "KaKs_Calculator 2.0: A Toolkit Incorporating Gamma-Series Methods and Sliding Window Strategies." *Genomics, Proteomics & Bioinformatics* 8, no. 1: 77–80.

Wang, Y., H. Tang, J. D. Debarry, et al. 2012. "*MCScanX*: A Toolkit for Detection and Evolutionary Analysis of Gene Synteny and Collinearity." *Nucleic Acids Research* 40, no. 7: e49.

Wang, Y.-J., T. Tain, J. Y. Yu, et al. 2023. "Genomic and Structural Basis for Evolution of Tropane Alkaloid Biosynthesis." *Proceedings of the National Academy of Sciences of the United States of America* 120, no. 17: e2302448120.

Wang, Z., N. Hobson, L. Galindo, et al. 2012. "The Genome of Flax (*Linum usitatissimum*) Assembled De Novo From Short Shotgun Sequence Reads." *Plant Journal* 72, no. 3: 461–473.

Wen, B., and S. Song. 2005. "A Brief Research Into the Rare and Endangered Plants in Xishuangbanna." *Journal of Central South Forestry University* 25, no. 2: 50–54.

World Conservation Monitoring Centre. 1998. "Pellacalyx yunnanensis." *The IUCN Red List of Threatened Species 1998*: e.T32436A9706845.

Wu, S., B. Han, and Y. Jiao. 2019. "Genetic Contribution of Paleopolyploidy to Adaptive Evolution in Angiosperms." *Molecular Plant* 13, no. 1: 59–71.

Wu, Z., and P. H. Raven. 1994. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.

Xu, S., Z. He, Z. Zhang, et al. 2017. "The Origin, Diversification and Adaptation of a Major Mangrove Clade (Rhizophoreae) Revealed by Whole-Genome Sequencing." *National Science Review* 4, no. 5: 721–734.

Xu, S., S. Shao, X. Feng, et al. 2024. "Adaptation in Unstable Environments and Global Gene Losses—Small but Stable Gene Networks by the May-Wigner Theory." *Molecular Biology and Evolution* 41, no. 4: msae059.

Xu, Z., and G. Tao. 1987. "Discussion on the Method of Systematic Assessment to Regional Threatened Plants and Their Prior Conservation." *Acta Botanica Yunnanica* 9, no. 2: 192–202.

Yang, F., L. Cai, Z. Dao, and W. Sun. 2022. "Genomic Data Reveals Population Genetic and Demographic History of *Magnolia fistulosa* (Magnoliaceae), a Plant Species With Extremely Small Populations in Yunnan Province, China." *Frontiers in Plant Science* 13: 1–13.

Yang, J., L. Cai, D. Liu, G. Chen, J. Gratzfeld, and W. Sun. 2020. "China's Conservation Program on Plant Species With Extremely Small Populations (PSESP): Progress and Perspectives." *Biological Conservation* 244: 108535.

Yang, Y., T. Ma, Z. Wang, et al. 2018. "Genomic Effects of Population Collapse in a Critically Endangered Ironwood Tree *Ostrya rehderiana*." *Nature Communications* 9, no. 1: 1–9.

Yang, Z. 2007. "PAML 4: Phylogenetic Analysis by Maximum Likelihood." *Molecular Biology and Evolution* 24, no. 8: 1586–1591.

Yu, G., L. G. Wang, Y. Han, and Q. Y. He. 2012. "clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters." *OMICS: A Journal of Integrative Biology* 16, no. 5: 284–287.

Yunnan Bureau of Geology and Mineral Resources. 1995. *Atlas of the Sedimentary Facies and Palaeogeography of Yunnan*. Kunming: Yunnan Science and Technology Press.

Zachos, J., H. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. "Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present." *Science* 292, no. 5517: 686–693.

Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. "An Early Cenozoic Perspective on Greenhouse Warming and Carbon-Cycle Dynamics." *Nature* 451, no. 7176: 279–283.

Zhang, X., G. Wang, S. Zhang, et al. 2020. "Genomes of the Banyan Tree and Pollinator Wasp Provide Insights Into Fig-Wasp Coevolution." *Cell* 183, no. 4: 875–889.e17.

Zheng, B., Q. Xu, and Y. Shen. 2002. "The Relationship Between Climate Change and Quaternary Glacial Cycles on the Qinghai-Tibetan Plateau: Review and Speculation." *Quaternary International* 97–98: 93–101.

Zhou, R., D. Macaya-Sanz, C. H. Carlson, et al. 2020. "A Willow Sex Chromosome Reveals Convergent Evolution of Complex Palindromic Repeats." *Genome Biology* 21, no. 1: 1–19.

Zhu, H. 2022. "Tropical Rain Forest of Yunnan (Southwestern China): Characteristics, Biogeographical Origin and Evolution." *Journal of Tropical and Subtropical Botany* 30, no. 4: 575–591.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.