scientific reports

Check for updates

OPEN

Reduced representation sequencing reveals genetic diversity and adaptive genetic divergence in *Calamus rhabdocladus*

Rui Gu[1,2,5], Ruijing Xu[1,5], Songpo Wei[1], Pär K. Ingvarsson[4], Shaohui Fan[1,2] & Guanglu Liu[1,3]✉

Small, isolated populations are more susceptible to natural disturbances and declines in genetic diversity. *Calamus rhabdocladus* has a scattered and fragmented natural distribution across South China, where only a few relict populations still exist. However, no population genetic studies have been conducted to date, and conservation management plans are lacking. In this study, we examined the genetic diversity, genetic differentiation, and environmental adaptability of all sampled individuals of *C. rhabdocladus* using reduced representation sequencing technology. We observed a moderate level of genetic diversity, with the values of unbiased observed heterozygosity (Ho), unbiased expected heterozygosity (He), Nei's gene diversity (Nei's), and Shannon's information index (I) being 0.180, 0.282, 0.279, and 0.436, respectively. Additionally, the Guangzhou (GZ), Heyuan (HY), and Foshan (FS) sites exhibited higher genetic diversity than other sites. Large genetic differentiation was detected in *C. rhabdocladus*. Structure analysis, phylogenetic tree, and principal component analysis all clustered the samples into three distinct genetic groups, shaped by geographical barriers, heterogeneous environments, and germplasm reproduction characteristics. Redundancy analysis (RDA) identified the driest month (bio14) and the mean temperature of the wettest quarter (bio8) as the two most important environmental factors driving adaptive genetic variation. The Latent Factor Mixed Model (LFMM) identified 955 loci significantly associated with environmental adaptation and successfully aligned 469 genes from related species, which were enriched in functional categories involving cell structure, carbohydrate metabolism, and protein degradation. Overall, this study represents the first investigation into the genetic diversity and structure of *C. rhabdocladus*. Our findings will aid in the conservation and utilization of the species, providing a theoretical basis for future ecological and evolutionary research.

Keywords *Calamus rhabdocladus*, Genetic variation, Environmental adaptation, Geographical isolation, Conservation strategy

Genetic diversity is crucial for the long-term survival of species, as it supports their evolutionary potential in response to environmental fluctuations[1]. Several factors influence genetic diversity, including population size, geographic range, reproductive patterns, and environmental variability[1,2]. Generally, species with smaller populations and restricted ranges exhibit lower genetic diversity than those with larger populations and broader distributions[3]. Small populations are particularly vulnerable to genetic drift, which can further decrease genetic diversity and result in significant genetic differentiation over time[4,5]. A reduction in genetic diversity may impair population resilience, limit the species' evolutionary potential, increase the risk of inbreeding depression, and ultimately raise the likelihood of local extinction[6].

[1]International Center for Bamboo and Rattan, Beijing, China. [2]Key Laboratory of National Forestry and Grassland Administration Beijing for bamboo & Rattan Science and Technology, Beijing, China. [3]Yunnan Diannan Bamboo Forest Ecosystem Research Station, Yunnan, China. [4]Department of Plant Biology, Linnean Centre for Plant Biology, Swedish University of Agricultural Sciences, Uppsala, Sweden. [5]These authors contributed equally: Rui Gu and Ruijing Xu. ✉email: liuguanglu@icbr.ac.cn

Calamus rhabdocladus Burret (Arecaceae family, subfamily Calamoideae, and genus *Calamus*) is a rare climbing plant mainly distributed in tropical and subtropical forests. In China, it is sporadically found in the southwestern regions, such as Hainan, Guangxi, and Guangdong provinces, growing on slopes and near river valleys[7]. *Calamus rhabdocladus* is a valuable non-timber multifunctional forest resource, ranking just below wood and bamboo as a forest product[8]. Owing to its distinct features, such as softness and resistance to bending and tension, *C. rhabdocladus* has become a natural raw material for construction, weaving, furniture, and handicrafts, earning it the nickname "green gold"[9]. The worldwide demand for both unprocessed and processed canes is substantial, with the total export trade value of rattan products currently being $3.5 billion[8]. Due to multiple threats, including habitat destruction, overexploitation, climate change, and poor regenerative ability, the population size and number of *C. rhabdocladus* individuals have sharply decreased[9], leading to a significant loss of genetic resources. Furthermore, low seed production, difficulties in seedling regeneration, and the lack of an effective long-distance seed dispersal mechanism have hindered population recovery[7]. Additionally, in the artificial cultivation of *C. rhabdocladus*, productivity and quality are constrained by limitations in germplasm resources, site conditions, cultivation methods, and harvest cycles[8]. As a result, the remaining wild populations are fragmented, small, and dispersed. Therefore, there is an urgent need to conserve the genetic resources of this vulnerable species, as understanding their genetic background is a prerequisite for implementing effective conservation strategies. However, current studies on *C. rhabdocladus* have primarily focused on its cultivation methods, ecological characteristics, and vine properties[7,10], no research has yet been conducted on its genetic diversity and differentiation.

A phenotypic study of *C. rhabdocladus* across eight different geographical regions with varying climatic conditions revealed significant morphological diversity, classified the germplasm into four distinct groups and provided evidence for local adaptation[11]. In plants, a substantial portion of phenotypic variation can be attributed to divergent selection driven by environmental variables[12,13]. Varying environmental conditions foster unique ecological communities and genetic structure, driving plants to develop favorable traits that support their growth, resulting in a diverse array of phenotypes across environments[14,15]. However, the key environmental variables that drive selection among populations remain unknown for most plants, and in line with this, the genetic mechanisms underlying adaptive divergence in *C. rhabdocladus* have yet to be explored in any greater detail.

Single nucleotide polymorphisms (SNPs) are the most common and widespread type of genetic variation in DNA sequences and these genetic markers have important applications in genetic diversity research[16,17]. With advances in high-throughput sequencing, genome-wide SNP detection has become a powerful tool for identifying species-differentiating genetic variation and assessing its functional consequences[18]. For non-model species like *C. rhabdocladus*, restriction site-associated DNA sequencing (RAD-seq) offers a cost-effective, reference-free, and high-throughput approach to resolving genetic variation, making it widely used in population genetics and evolutionary studies[19–22]. So far, no studies have utilized the RAD-seq approach to explore the genetic diversity and structure of *C. rhabdocladus*. We hypothesized that the genetic diversity of *C. rhabdocladus* would be low due to the species' limited population size and fragmented distribution, with biogeographic barriers and heterogeneous habitats further driving genetic differentiation. In this study, we employed RAD-seq technology to investigate genetic diversity and differentiation across 10 natural populations spanning the species' entire Chinese distribution. Our aims were to: (1) uncover the genetic diversity and structure of *C. rhabdocladus*; (2) determine the key factors shaping genetic variation; and (3) identify the genetic mechanisms underlying adaptation to different habitat conditions. Based on our findings, we provide specific suggestions for the further protection and sustainable use of wild *C. rhabdocladus* resources.

Results

SNPs discovery

A collection of 47 *C. rhabdocladus* individuals was effectively sequenced using the HiSeq Nova-600 system (Illumina Inc., USA), generating 890.92 million raw reads. Post-quality filtering yielded 868.17 million high-quality reads, averaging 18.47 million reads per sample. The data quality was exceptional, with Q20 values between 96.77% and 98.14%, and Q30 values from 91.45 to 94.50%. The GC content was consistent, ranging from 39.21 to 41.95% (Table 1; Table S1). Initially, 10,644,342 SNPs were detected across all the samples. Under stringent filtering, 2,001,229 high-quality SNPs were retained for further analysis. The majority of SNPs were base transitions (70.85%), with the proportions of A/G, T/C, and A/C transitions being 35.39%, 35.86%, and 10.02%, respectively (Fig. S1; Table S2).

Genetic diversity

C. rhabdocladus displayed moderate genetic diversity at the species level, with values of unbiased observed heterozygosity (Ho), unbiased expected heterozygosity (He), Nei's gene diversity (Nei's), and Shannon's

	Clean reads (bp)	Clean bases (bp)	GC(%)	Q20(%)	Q30(%)
total	868,174,594	130,210,138,184	-	-	-
max	22,942,380	3,441,040,314	41.95	98.14	94.5
min	14,177,824	2,126,388,270	39.21	96.77	91.45
avg	18,471,799	2,770,428,472	40.29	97.25	92.35

Table 1. RAD-seq data statistics of 47 *C. rhabdocladus* individuals.

information index (I) being 0.180, 0.282, 0.279, and 0.436, respectively (Table 2). Meanwhile, among the *C. rhabdocladus* sampling sites, the GZ site exhibited the highest genetic diversity (He = 0.213, Ho = 0.243, Nei's = 0.191, I = 0.254), followed closely by the HY (He = 0.210, Ho = 0.237, Nei's = 0.188, I = 0.251) and FS sites (He = 0.213, Ho = 0.247, Nei's = 0.185, I = 0.238). In contrast, the ZZ site displayed the lowest genetic diversity, with He and Ho of 0.188 and 0.2, respectively. The estimated inbreeding coefficient (Fis) ranged from −0.264 at the LB site to 0.042 at the NN site. Due to the excess of observed heterozygotes, Fis is negative for all sites except NN, indicating low levels of inbreeding.

Genetic structure

An analysis of the genetic structure was conducted to determine the optimal number of genetic groups explaining the observed variation. The lowest cross-validation (CV) error was obtained when K = 2, which separated all samples into two groups (Fig. 1). However, some individuals from the NN and LB sites showed genetic admixture. The second-lowest CV error was observed at K = 3, revealing a clear genetic structure with minimal admixture. Under K = 3, the ten sites were divided into three clusters: the SY, CJ, and LD sites formed a southern cluster (group I) on Hainan Island, the LB and NN sites constituted a central cluster (group II), and the remaining sites (ZZ, HY, GZ, FS, and CZ) formed a larger cluster (group III), which spans across the Nanling Mountains (Fig. 2). The results of the PCA and maximum likelihood (ML) tree analysis for the 47 individuals aligned with those of the structure analysis. In the PCA, the first three principal components explained approximately 58% of the total variation, dividing the individuals into three primary groups (Fig. 3). Individuals from ZZ, HY, GZ, FS, and CZ were evenly distributed along PC1 and PC2. The NN and LB sites were positioned intermediately and were markedly differentiated on PC3, while the CJ, SY, and LD sites were located at the opposing extreme on PC1. The ML tree also displayed similar genetic groupings (Fig. 4), supporting the reliability of the genetic clusters. The inferred genetic clusters corresponded with the geographic barriers between sites (see Results 3.4). Mantel tests revealed a significant positive association between genetic and geographic distance ($r = 0.687$, $P < 0.001$), indicating strong isolation by distance (IBD). Due to this strong pattern of genetic IBD, we conducted a cross-validation analysis using both spatial and non-spatial models in conStruct, which revealed that the spatial model had superior predictive accuracy compared to the non-spatial model (Figs. S2 and S3). When K reached 3, the predictive accuracy leveled off with minimal further increase. In this model, spatial analyses divided the total genetic variance into three clusters, suggesting that the divergence may be due to dispersal barriers rather than solely isolation by distance (IBD).

Genetic differentiation, gene flow and genetic barrier

The AMOVA analysis showed that 80.43% of the total genetic variance could be attributed to individuals within sampling sites, compared to 19.57% among sampling sites ($P < 0.01$) (Table 3). At the group level, 95.65% of the total genetic variance was due to differences within groups, with only 4.35% occurring among groups ($P < 0.01$). Pairwise Fst values ranged from 0.069 (between HY and GZ) to 0.572 (between CJ and ZZ), with an average of 0.341 (Table S3). Additionally, among the three groups, the highest Fst value was observed between group I and group II (0.442), followed by group I and group III (0.393), with the lowest between group II and group III (0.226). This suggests that geographic distance between groups may directly influence the pairwise Fst values. Gene flow (Nm) values varied from 0.225 (between CJ and ZZ) to 3.38 (between HY and GZ), averaging 0.798, indicating limited genetic exchange among *C. rhabdocladus* sites (Table S4). A genetic barrier analysis identified two likely barriers to gene flow. The first barrier (a), with strong bootstrap support (65–70%), isolated group

Population code	Population name	Lat.(*N*)/Long.(E)	Altitude	Number of individuals	Cluster	Ho	He	Nei's	I	Fis
SY	Sanya, Hainan	18.37/109.66	258.6	5	I	0.209	0.2	0.181	0.239	−0.063
CJ	Changjiang, Hainan	19.09/109.19	1124	5	I	0.214	0.2	0.18	0.238	−0.091
LD	Longdong, Hainan	18.71/108.88	880.2	5	I	0.215	0.2	0.181	0.24	−0.091
group I						0.188	0.274	0.275	0.395	-
LB	Libo, Guizhou	25.48/108.24	383	3	II	0.263	0.209	0.17	0.205	−0.264
NN	Nanning, Guangxi	23.56/108.35	598.4	5	II	0.196	0.211	0.181	0.251	0.042
group II						0.2	0.308	0.276	0.425	-
CZ	Chenzhou, Hunan	25.38/113.66	327.8	5	III	0.219	0.201	0.181	0.24	−0.103
FS	Foshan, Guangdong	22.78/112.44	143.3	4	III	0.247	0.213	0.185	0.238	−0.169
GZ	Guangzhou, Guangdong	23.76/113.90	266.6	5	III	0.243	0.213	0.191	0.254	−0.15
HY	Heyuan, Guangdong	23.47/114.75	126.4	5	III	0.237	0.21	0.188	0.251	−0.126
ZZ	Zhangzhou, Guangdong	24.51/117.24	336	5	III	0.2	0.188	0.171	0.225	−0.081
group III		-				0.186	0.286	0.285	0.427	-
Mean		-				0.224	0.205	0.181	0.238	-
Species level		-		47		0.18	0.282	0.279	0.436	-

Table 2. Geographic locations and genetic diversity in ten populations of *C. Rhabdocladus*. *Ho* unbiased observed heterozygosity; *He* unbiased expected heterozygosity; *Nei's* Nei's gene diversity; *I* Shannon's information index; *Fis* inbreeding coefficient.

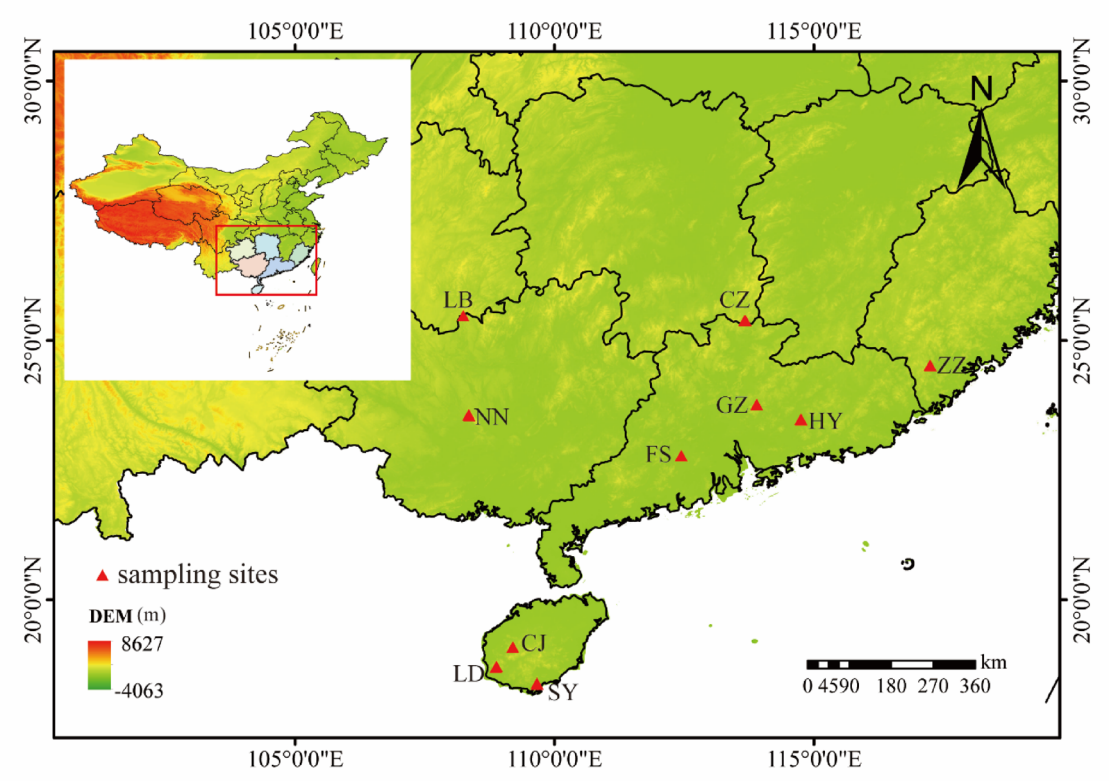

Fig. 1. Genetic structure analysis of 47 individuals of *C. rhabdocladus* with K values ranging from 2 to 3. (**a**) The plot of cross-validation error with changing K values; (**b**) Each vertical bar represents the estimated posterior probability of each individual in the genetic clusters (K = 2, 3). The name of each individual is shown below the bar plot.

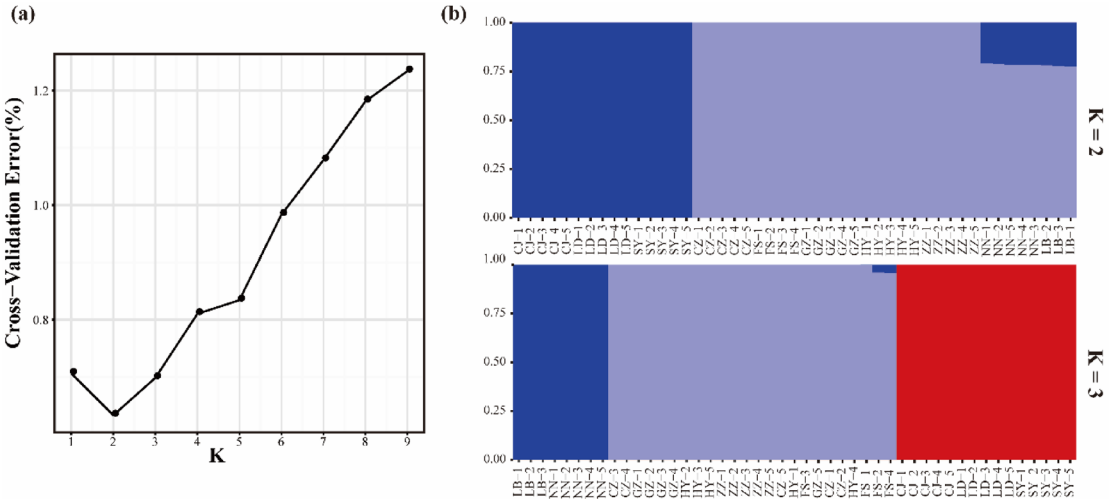

Fig. 2. Location of the sampling site in China.

I from the other two groups. The second, weaker barrier (b), with 48–54% bootstrap support, was detected between group II and group III (Fig. 5).

Association between genetic variation and environmental factors

Mantel tests revealed a strong positive correlation between genetic and environmental distances ($r = 0.600$, $P < 0.001$). To examine the effect of environmental factors on genetic variation, we conducted a redundancy analysis (RDA) (Table 4; Fig. 6). The analysis showed that the first two RDA axes captured a substantial portion of the genetic variation (33.44% by RDA1 and 13.76% by RDA2), with both axes demonstrating a strong correlation

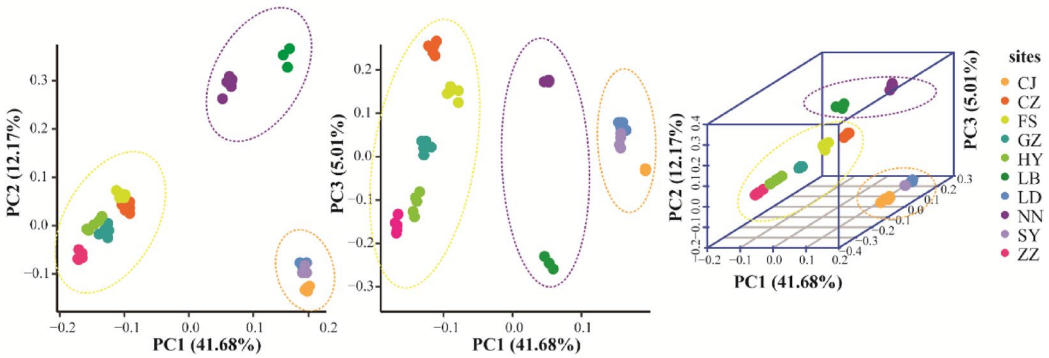

Fig. 3. Principal component analysis (PCA) of the 47 *C. rhabdocladus* individuals based on the first three principal components. The yellow dashed ellipse represents group I, the purple dashed ellipse represents group II, and the orange dashed ellipse represents group III.

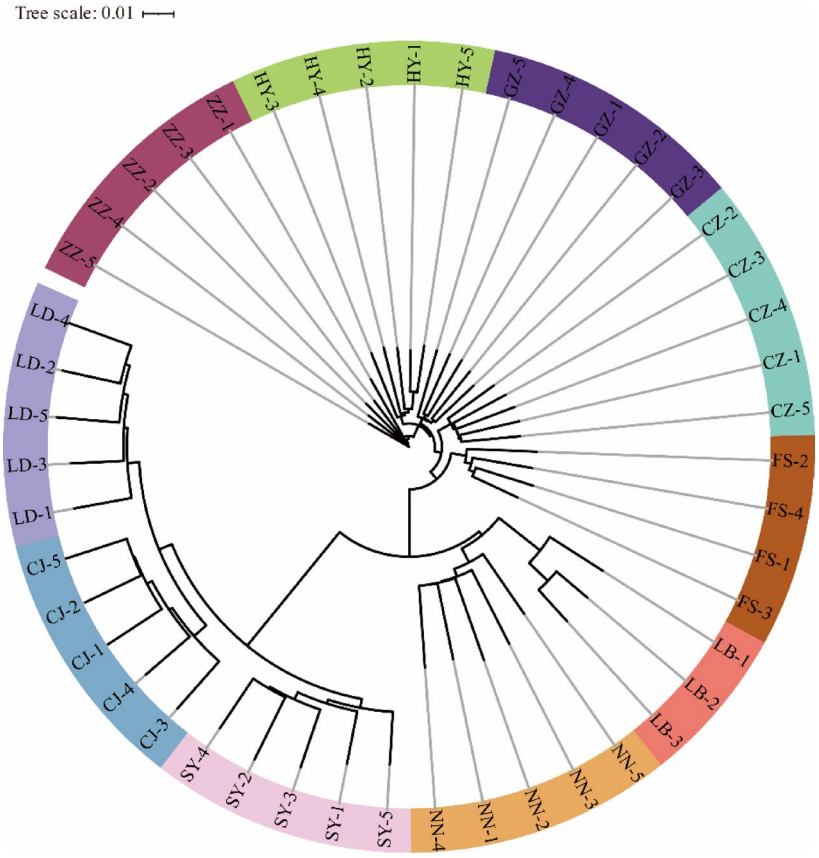

Fig. 4. Maximum likelihood (ML) tree based on SNPs of 47 *C. rhabdocladus* individuals. Different colors represent various sampling sites.

(0.95) with gene frequencies across the ten environmental variables. These variables effectively grouped the ten sites into three clusters, aligning with the population structure analysis. Notably, seven environmental variables showed significant correlations with the first two RDA axes (Table 4; Fig. 6), highlighting their influence on genetic variation. Among these, four variables—minimum temperature of the coldest month (bio6), temperature annual range (bio7), mean temperature of the wettest quarter (bio8), and mean temperature of the driest quarter (bio10)—were related to temperature, while three variables—precipitation of the driest month (bio14), precipitation seasonality (bio15), and precipitation of the warmest quarter (bio18)—were related to precipitation. Bio14 and Bio8 were identified as the two key environmental drivers that contribute to the genetic variation within sites (Fig. 6; Table S5).

Spatial scale	Source of variation	Degree of freedom	Sum of squares	Mean square	Variance components	Percentage of variance (%)	*P*-value
Sites	Among sites	9	1.226	0.136	0.028	19.57	< 0.01
	Within sites	37	0.198	0.005	0.005	80.43	< 0.01
	Total	46	1.424	0.031		100	
Groups	Among groups	2	0.035	0.017	−0.001	4.35	< 0.01
	Within groups	44	1.389	0.032	0.032	95.65	< 0.01
	Total	46	1.424	0.031		100	

Table 3. Analysis of molecular variance (AMOVA) within/among *C. rhabdocladus* sampling sites/groups.

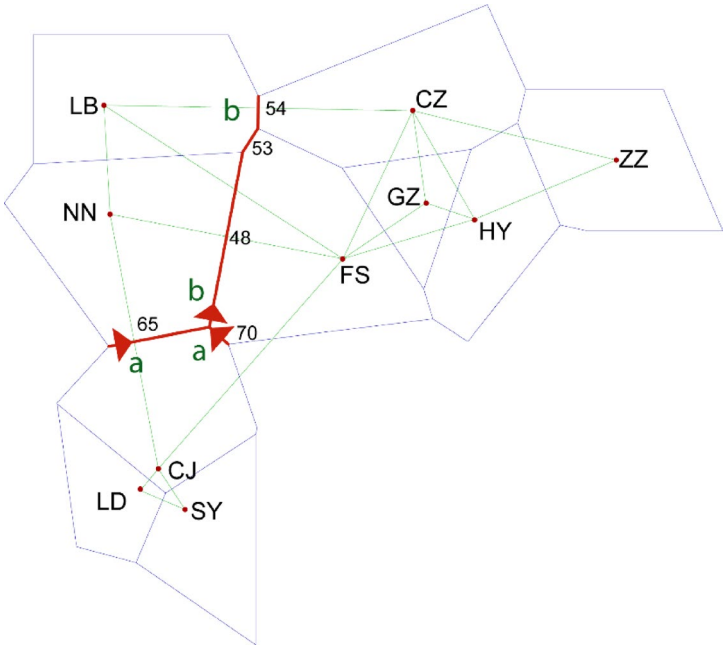

Fig. 5. Predicted genetic barriers to gene flow. Blue lines correspond to hypothetical boundaries between sites, labeled with corresponding codes. Red lines with arrows indicate barriers to gene flow.

Environmental variable	Axis 1	Axis 2	Axis 3	Axis 4
bio1	−0.184	−0.056	−0.115	−0.142
bio6	−0.5768**	0.025	0.052	0.012
bio7	0.722**	−0.183	0.032	−0.164
bio8	0.002	0.365*	−0.146	−0.295*
bio10	0.385**	−0.058	−0.331*	−0.603**
bio13	0.125	0.002	0.141	0.025
bio14	0.576**	0.17	−0.516**	−0.214
bio15	−0.757**	0.315*	0.088	0.112
bio16	0.167	0.087	−0.003	0.047
bio18	0.201	0.352*	0.029	0.222

Table 4. Associations between environmental factors and the axes of ordination. Statistically significant correlation by *($P < 0.05$) and **($P < 0.01$).

Identification of local adaptation loci and annotation

To detect genomic loci associated with local adaptation, we performed LFMM analyses to test correlations between single loci and single environmental variables. The analysis identified 955 SNPs significantly correlated with three environmental variables: bio14 (865 SNPs), bio8 (82 SNPs), and bio18 (9 SNPs). Since LFMM is a univariate detection method, we also conducted a genotype-environment association analysis using multivariate landscape genomic method, redundancy analysis (RDA). This analysis further pinpointed 1,563 SNPs with

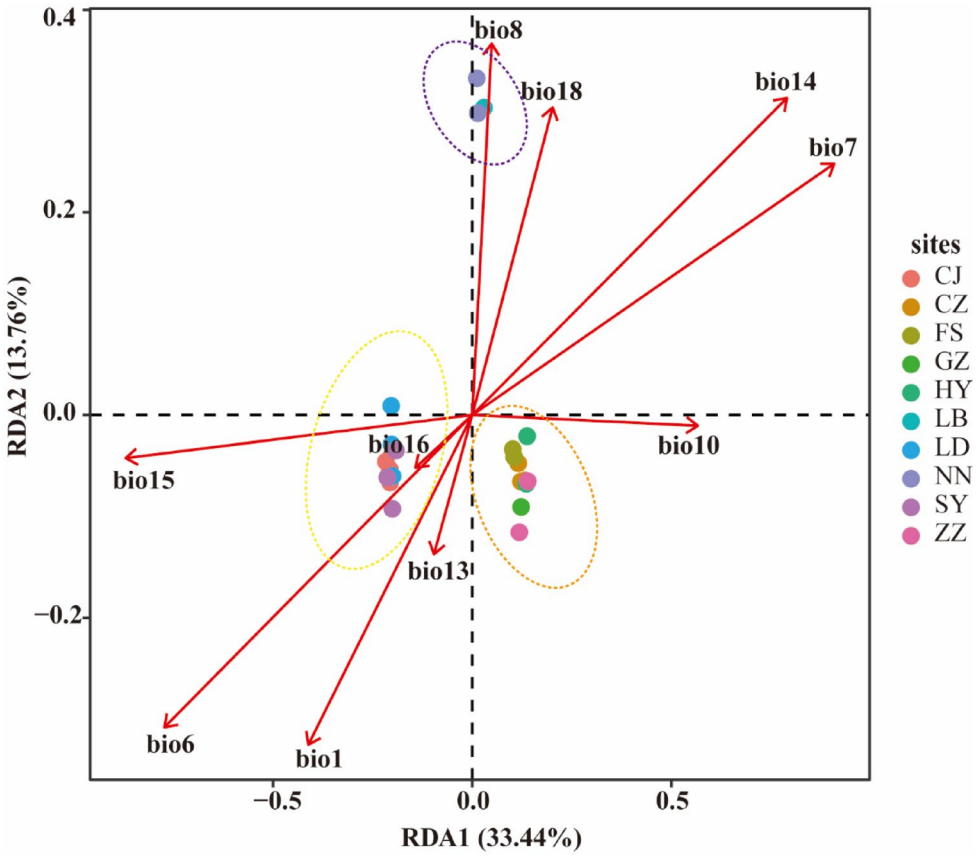

Fig. 6. Redundancy analysis (RDA) of *C. rhabdocladus* showing the relationship between genomic variation and ten environmental variables. Different colored dots represent different sampling sites. The yellow dashed ellipse represents group I, the purple dashed ellipse represents group II, and the orange dashed ellipse represents group III.

environmental associations, including a subset of 6 SNPs that were also identified by LFMM. An alignment of environment-associated loci with genomic information from a related species revealed that a total of 585 loci could be successfully matched to 469 genes, with 366 and 107 of these genes having annotation in the GO and KEGG databases, respectively. Among these genes, 425, 48, and 5 were associated with bio14, bio8, and bio18, respectively. A GO enrichment analysis indicated that the genes related to bio14 were classified into three main categories: 177 genes in "Cellular Component," 285 genes in "Molecular Function," and 256 genes in "Biological Process" (Table S6). The GO enrichment analysis of the 425 genes showed significant enrichment (q-values <0.05) in "cytoplasm" (GO:0005737), "Golgi membrane" (GO:0000139), "Golgi apparatus subcompartment" (GO:0098791), "cellular carbohydrate metabolic process" (GO:0044262), "trans-Golgi network" (GO:0005802), and "organelle subcompartment" (GO:0031984) (Tables S7 and S8). These genes may help plants adapt to drought by regulating mechanisms such as water utilization, metabolic adaptation, structural protection, and signal transduction. The genes associated with bio8 were classified into three main categories: 24 genes in "Cellular Component," 31 genes in "Molecular Function," and 31 genes in "Biological Process" (Table S9). These genes were significantly enriched in only two GO terms: "mycotoxin metabolic process" (GO:0043385) and "mycotoxin biosynthetic process" (GO:00433856), and one KEGG pathway, "porphyrin metabolism" (ko00860) (Tables S10 and S11). These genes may enhance plant disease resistance by regulating mycotoxin metabolic pathways to boost defense capabilities. The genes associated with bio18 were significantly enriched in 14 GO terms, mainly including "alpha-1,4-glucosidase activity" (GO:0004558), "maltose alpha-glucosidase activity" (GO:0032450), and "threonine-type endopeptidase activity" (GO:0004298) in "Molecular Function," "proteasome core complex" (GO:0005839), "endopeptidase complex" (GO:1905369) in "Cellular Component." One enriched KEGG pathway, "proteasome" (K02727), was also detected (Tables S12, S13 and S14). These genes play important roles in regulating carbohydrate and protein metabolism and maintaining cellular homeostasis, helping the plant maintain normal physiological functions under varying precipitation conditions in the warmest season.

Discussion

The availability of genomic data provides valuable opportunities for attaining a comprehensive understanding of evolutionary mechanisms affecting a species. Considering the absence of a reference genome, we used a simplified genome sequencing method to generate genomic data for *C. rhabdocladus*. This study is the first comprehensive

analysis of its genetic information. However, due to limitations in sample size at certain collection sites, we present the following insights based on our data.

Maintaining genetic diversity is essential for a species to evolve and adapt to changing environments[1]. Threatened species with relatively small population sizes and limited distribution ranges generally exhibit low genetic diversity[23,24]. However, in the study, despite the small number of individuals at each site, *C. rhabdocladus* demonstrated moderate genetic diversity at the species level, which contradicts our initial hypothesis. The Nei's genetic diversity index (Nei's = 0.279) and Shannon's diversity index (I = 0.436) for *C. rhabdocladus* were slightly higher than those observed in some wild rattans, such as *Daemonorops margaritae* (Nei's = 0.258, I = 0.388), *C. simplicifolius* (Nei's = 0.223, I = 0.344)[25], *Calamus flagellum* (Nei's = 0.233, I = 0.369)[26], *Calamus palustris* var. *malaccensis* (Nei's = 0.153)[27], *Calamus thwaitesii* (Nei's = 0.272–0.333)[28], and *Calamus vattayila* (Nei's = 0.150)[29]. The genetic diversity of *C. rhabdocladus* is also consistent with levels reported for ten other palm species[30]. In contrast, the expected average heterozygosity (He = 0.282) of *C. rhabdocladus* was lower than that of other rare and endemic tree species, such as *Boswellia papyrifera* (Del. ex Caill.) Hochst (He = 0.69)[31], *Dalbergia cochinchinensis* (He = 0.55)[32], and *Ottelia acuminata* (He = 0.35)[33]. The observed moderate level of genetic diversity in this study may be linked to the outcrossing nature of *C. rhabdocladus*[7]. Previous studies suggest that outcrossing plants generally exhibit greater genetic diversity than self-compatible species[34,35]. Additionally, the long lifespan of *C. rhabdocladus* allows for the accumulation of genetic mutations over time, thereby enhancing genetic diversity[7,36]. While our study revealed moderate genetic diversity in *C. rhabdocladus*, we acknowledged these findings should be interpreted cautiously as conventional metrics rely on neutral theory assumptions - a framework increasingly challenged by genomic evidence[37,38], whereas emerging paradigms like the maximum genetic diversity theory may better explain diversity preservation in bottlenecked populations. Interestingly, at the site level, the observed heterozygosity is higher than the expected heterozygosity, whereas at the group level, it was lower than expected, that is likely due to the Wahlund effect (i.e. due to substructure within groups). We observed that genetic diversity parameters (I, Ho, He, Nei's) varied across different sites. The GZ, HY, and FS sites exhibited higher genetic diversity compared to other sites, likely due to their locations near the coast and the Guangdong Mountain range. The unique terrain and climatic conditions in these areas may have provided refuge for the species and suggests that species might have had a longer evolutionary history in this region. This also suggests that these sites could have served as a source of useful and rare genetic variation in the past, and conservation efforts should therefore be intensified to protect the genetically diverse specimens present in these locations.

The genetic structure of species reflects the spatial and temporal patterns of genetic variability, which may be influenced by geographical or ecological barriers, the species' life history, geological events, and selective pressures[39]. In this study, we used three methods to assess the genetic diversities and genetic relationships of the 47 accessions, ensuring reliable results. The maximum likelihood tree grouped all individuals into three clusters, indicating three genetically distinct groups. Structure analysis also identified these three clusters, with minimal genetic admixture, reflecting limited gene flow between certain sites (supported by the low Nm estimates for certain site pairs). This division displayed a spatial pattern that aligns with their geographical locations. The AMOVA results ($P < 0.001$) further confirmed significant genetic differentiation among the groups.

What factors have shaped the present genetic structure of the species? The first potential reason may be the low efficiency of seed dispersal. Species with limited dispersal abilities typically exhibit a clear hierarchical genetic structure, whereas those that disperse widely are less likely to[40]. *C. rhabdocladus* is an insect-pollinated species that reproduces sexually through seeds, with each fruit containing only one to three seeds. However, the species has low fruiting rates in the wild[7]. Establishing populations in new suitable habitats requires multiple independent seed dispersal events, which significantly reduce the likelihood of long-distance seed dispersal[41]. This limited seed dispersal distance leads to relatively low gene flow (0.798) among groups, resulting in spatial isolation and the formation of distinct genetic groups. Furthermore, this limited distance may be insufficient to prevent genetic differentiation due to genetic drift, especially in small and isolated populations. Hence, genetic drift may also influence the genetic structure of *C. rhabdocladus*. The second potential factor contributing to the current spatial distribution is geographical barriers. Two genetic barriers were directly detected between the three groups. The first barrier, the Qiongzhou Strait, likely serves as a geographical boundary isolating group I from the other groups. The second barrier, the Nanling Mountains, may separate groups II and III. ConStruct analysis also supports the presence of overall isolation-by-distance (IBD). These discontinuous mountains and the strait further block gene flow, leading to genetic differentiation between the groups. The third possible reason for the genetic differentiation is environmental differences. Long-term geographical isolation and fragmentation may have caused these groups to occupy distinct ecological niches[15]. The AMOVA showed that a large proportion of the total variance was attributed to within group variance, suggesting that significant genetic variation may arise from local adaptation to heterogeneous environments within the groups despite geographic barriers. A Mantel test indicated a significant association between genetic and environmental distances, and an RDA analysis demonstrated that environmental variables could effectively subdivide *C. rhabdocladus* into three groups, corroborating the results from the populations structure analyses. The RDA analysis also identified bio14 and bio8 as the two most crucial variables associated with genetic variation. Environmental heterogeneity may lead to variation in allele frequencies of genes under natural selection, resulting in enhanced genetic differentiation among the three groups at such sites. Group I experienced higher minimum temperatures in the coldest month, higher mean temperatures in the wettest quarter, and greater precipitation in the driest month, while group III was subject to a larger annual temperature range. Finally, the role of historical population events, such as migrations, expansions, and bottlenecks, in shaping the current genetic structure warrants further investigation.

Adaptation to local environmental conditions significantly contributes to the structuring of genetic variation. By aligning adaptive loci with the genome annotation information of the related species *C. simplicifolius*, we identified several key candidate genes for local adaptation. Among these, the candidate genes associated

with bio14 were significantly enriched in biological processes such as cell wall synthesis and modification, carbohydrate metabolism, signal transduction, plant resistance, and water balance regulation (q-value < 0.05). For instance, the AT-hook motif nuclear localized (AHL) transcription factor encoded by *DY01Gene08175* enhances plant drought tolerance by regulating the expression of genes related to carbohydrate metabolism and coordinating the plant's physiological and metabolic responses[42]. *DY01Gene16368*, encoding Magnesium/Proton Exchanger 1 (MHX1), and *DY01Gene26241*, encoding Metal Tolerance Protein 8 (MTP8), maintain ion balance by exchanging and transporting metal ions and protons[43]. *DY01Gene34175* encodes Palmitoyl-protein thioesterase 1 (PPT1), which regulates protein functions through depalmitoylation. These proteins might be downstream effectors of the transcription factor CkWRKY33, which is involved in *Arabidopsis*'s responses to environmental stresses such as drought and salt stress, indirectly influencing drought response[15]. Overall, the synergistic action of these genes helps *C. rhabdocladus* maintain physiological balance and adapt to drought conditions. Genes associated with bio8 were significantly enriched in mycotoxin metabolic and biosynthetic processes (q-value < 0.05). These genes might enhance plant defense responses, helping plants resist fungal infection and toxin damage in hot and humid environments. Additionally, the homolog of *DY01Gene33916* (At3g21340) is involved in regulating the salicylic acid (SA)-mediated immunity pathway and plays a significant role in modulating Arabidopsis's response to environmental stresses such as pathogen infection[44]. KEGG enrichment analysis revealed that *DY01Gene47010* and *DY01Gene32474* are significantly enriched in the porphyrin metabolism pathway (q-value < 0.05). The HY2 and PAO proteins encoded by these genes regulate pigment synthesis and degradation, helping plants maintain photosynthetic efficiency and adaptability in humid conditions. Genes associated with bio18 were significantly enriched in the ubiquitin-mediated protein degradation KEGG pathway. The ubiquitin-mediated protein degradation process influences various aspects of plant activities, such as hormone signal transduction, reproductive processes, and responses to abiotic stresses. These factors are crucial for the local adaptation of the *C. rhabdocladus* species[45]. Future research should focus on exploring the specific functions, interactions, regulatory networks, and pathways of more adaptive genes to further illuminate their roles in environmental adaptation. This will help in formulating effective conservation strategies for *C. rhabdocladus*.

From the perspective of resource conservation, small sampling sites are at high risk of extinction due to uncontrolled human exploitation, habitat degradation, and environmental fluctuations. Any reduction in individual numbers can substantially weaken the species' long-term resilience to changing environmental conditions. It is necessary to protect all individual plants in their natural habitats to maintain as much genetic variation as possible. Meanwhile, the GZ, HY, and FS sites, which represent relatively diverse genetic resources, should be prioritized in conservation efforts. Despite multiple investigations, only a few individuals were found in the LB and FS regions. To mitigate the risk of genetic decline, it is advisable to carry out artificial outcrossing and propagation in these locations. Furthermore, since these areas are not covered by any nature reserves, establishing mini-reserves is recommended to urgently protect the LB and FS populations. Besides in situ conservation efforts, establishing gene collection orchards for *C. rhabdocladus* is crucial for preserving a wide variety of genotypes. Additionally, with the advancement of ex situ conservation technologies, cryopreservation of cells, tissues, or DNA has become a viable approach for preserving superior germplasm[8]. From the perspective of resource cultivation, it is imperative to deepen research into the propagation techniques of *C. rhabdocladus*, including seed and vegetative propagation methods (e.g., stem cuttings, layering, division), to enhance the spread of high-quality specimens. Intensifying research on introduction, domestication, and afforestation techniques is also crucial, with a focus on environmental variables, especially bio14 and bio8. From the perspective of resource utilization, the study shows that *C. rhabdocladus* holds considerable breeding potential for drought resistance, demonstrating good adaptability to changes in the precipitation of the driest month. Future efforts should focus on analyzing the genetic mechanisms that enable adaptation to these critical environmental factors.

Overall, our study is the first to examine the genetic diversity and structure of wild *C. rhabdocladus* populations at the genomic level. *C. rhabdocladus* has moderate genetic diversity, with the GZ, HY, and FS sites as key diversity centers that should be prioritized for conservation. Geographic barriers, environmental variation, and reproductive traits have shaped its genetic structure. Comprehensive in situ and ex situ conservation strategies are crucial for the species' long-term survival.

Materials and methods

Field sampling

We collected distribution information from the Chinese Virtual Herbarium (CVH), GBIF and Junqing Nong's work[11], for the following ten regions: Sanya, Ledong, Changjiang, Nanning, Zhangzhou, Heyuan, Guangzhou, Foshan, Chenzhou, and Libo (Fig. 2). Plants were identified based on morphological traits following the species' formal description[46,47]. A total of 47 leaf samples from wild individuals were collected from ten *C. rhabdocladus* sites, which are naturally sporadic across nearly the entire regional distribution in China, with appropriate collection permits obtained. Vouchers specimens of the collected leaf samples were deposited at the International Bamboo and Rattan Centre Herbarium with the deposition number 47, and were formally identified by Ruijing Xu. To maintain consistency in sample size, five individual plants, spaced at least 100 m apart, were collected from each site. In locations where fewer than five individuals were available, all present individuals were sampled. All selected individuals were reproductive adults. We recorded the latitude and longitude information for each individual and collected a small amount of tender, healthy, mature leaves. These leaves were immediately sealed in plastic bags and preserved at −80 °C for subsequent Deoxyribonucleic Acid (DNA) extraction and genotyping. Table 2 presents comprehensive details for each sample site.

DNA extraction

An improved method of hexadecyltrimethylammonium bromide (CTAB) was employed to extract whole genomic DNA from approximately 20 milligrams of dried silica gel-preserved leaf samples for each individual[48]. To prevent oxidation of the sample DNA, the CTAB extraction buffer was supplemented with 2% beta-mercaptoethanol. Subsequently, the DNA degradation levels of each sample were evaluated by 1% agarose gel electrophoresis using a Bio-Rad PowerPac Basic (Bio-Rad, USA), and DNA purity was assessed using a NanoDrop (ND-2000, Thermo Scientific). The extracted DNA was then used for subsequent sequencing using a RAD-seq approach.

Library preparation and sequencing

Because Restriction Site-Associated DNA Sequencing (RAD-seq) provides better resolution than cpDNA and Simple Sequence Repeats (SSR) methods, even with a relatively small number of samples, 3–5 individuals were selected from each sampling site for sequencing[18]. RAD libraries were created, indexed, and pooled according to the methods outlined by Baird et al.[19]. The sequencing process for the *C. rhabdocladus* genomic DNA is as follows: (1) Restriction digestion: Genomic DNA was enzymatically digested by the restriction enzyme *Escherichia coli* Restriction Endonuclease I (*EcoRI*). (2) Ligation of P1 adapters: The restriction fragments were ligated with P1 adapters, which consist of four parts: a forward amplification primer recognition sequence, a primer recognition site, a 4–8 bp barcode sequence for sample tracking, and an *EcoRI* restriction enzyme recognition site. (3) Fragment selection: The DNA fragments with different barcode P1 adapters were mixed and randomly sheared. The fragments were detected by agarose gel electrophoresis, and fragments of 300–500 bp were recovered. (4) Ligation of P2 adapters: The recovered fragments were end-repaired by adding an "A" base, then ligated with P2 adapters. (5) Sequencing: The DNA fragments containing P1 and P2 adapters were amplified by PCR and purified with magnetic beads to generate the final library. Sequencing was conducted on the Illumina HiSeq Nova 600 platform using the Illumina PE 100 strategy. All experimental procedures were conducted by Geniotech Bio-pharm Technology Co., Ltd. in Guangzhou, China.

RAD data production and quality control

Due to the presence of low-quality sequences, adapter sequences, and other unwanted elements in the raw data, the following filtering criteria were applied to obtain high-quality sequencing data: (1) Reads from the raw data that do not contain the 4–8 bp barcode information at the 5' end were removed[18]. (2) After trimming the 5' end barcode from the remaining reads we performed restriction site recognition filtering to remove reads that did not contain the restriction enzyme recognition site[18]. (3) Reads containing adapter sequences were also removed[18]. (4) Reads where bases with a quality score less than 15 constituted more than 40% of the read were also filtered out[18]. (5) Finally, reads with more than 10% of bases being "N" (unknown bases) were discarded. The filtered data were then subjected to statistical analysis for read count, total base count, GC content, and quality scores (Q20, Q30)[18].

SNP detection and filtering

SNP detection was conducted using Stacks v1.48 (https://catchenlab.life.illinois.edu/stacks/)[49], with parameters set to require a minimum coverage depth of 2 for forming sequences into stacks, allowing a maximum of 2 mismatches between bases within stacks, and permitting a maximum of 3 mismatches between loci when building the catalog for SNP detection[18]. The specific steps were as follows: (1) Filtered reads were input into the ustacks program, where a hashing algorithm first clustered reads that perfectly matched into stacks[18]. These stacks were then compared to form a series of putative loci, with SNPs at each locus detected using a maximum likelihood framework[18]. (2) The cstacks program compiled loci from all individuals generated by ustacks into a catalog, merging alleles to produce a series of consistent loci[18]. (3) The sstacks program matched the series of stacks generated by ustacks with the catalog compiled by cstacks, determining the allelic state at each locus for every individual[18]. (4) The populations program was used for genetic statistics, and the SNP results were written into an output file in VCF format[18,49]. Based on prior various analyses and SNP filtering results, we retained the following SNPs suitable for our dataset: (1) biallelic loci; (2) missing rate less than 0.5[50]; (3) minor allele frequency greater than 0.05[51]; and (4) heterozygosity ratio less than 0.8. Finally, the VCF file was converted into multiple formats using PGDSpider v2.1.1.2 (https://software.bioinformatics.unibe.ch/pgdspider/) for subsequent statistical analyses[52]. To assess whether a small sample size affects genetic estimates, we calculated pairwise genetic differentiation by randomly reducing the number of samples from five to two per site[53–55]. We found no notable differences in fixation index (Fst) values: Fst [$n = 4$] = 0.299 (95% CI: 0.287–0.301), Fst [$n = 3$] = 0.303 (95% CI: 0.280–0.323), and Fst [$n = 2$] = 0.284 (95% CI: 0.280–0.288). Despite the limitations of small sample sizes, the high number of observed SNPs may help compensate for this, to some extent representing the genetic variability of each locality.

Genetic diversity and genetic differentiation analysis

The populations program in Stacks software was employed to calculate genetic diversity parameters[49], such as unbiased expected heterozygosity (He)[56], unbiased observed heterozygosity (Ho), Shannon's Information Index (I), Nei's gene diversity (Nei's), and the inbreeding coefficient (Fis). Additionally, DnaSP v6 (http://www.ub.edu/dnasp/) software was used to calculate the genetic differentiation coefficient (Fst) and gene flow (Nm)[57]. Nei's gene diversity (Nei's) was determined as Nei = nHe/(n-1), where n represents the number of samples. Shannon's Information Index (I) was computed using the equation $I = -\sum(Pi \times \ln(Pi))$, where Pi denotes the frequency of the i-th allele. Gene flow (Nm) was determined as $Nm = (1-Fst)/4Fst$[58]. A hierarchical analysis of molecular variance (AMOVA) was performed using Arlequin v3.5 (https://cmpg.unibe.ch/software/arlequin35/), with significance determined through 10,000 random permutations, to evaluate genetic variation both within and among sites

or groups identified in the genetic structure analyses *C. rhabdocladus*[59]. We used BARRIER v2.2(https://ecoanthropologie.fr/fr/logiciel-danalyse-geographique-barrier-22-9516) software with Monmonier's maximum difference algorithm to identify genetic barriers to gene flow between sites. This program mapped the sites using geographical coordinates and marked the barriers by assessing the highest values in the genetic distance matrix between the sampling points[60].

Genetic structure analysis

We performed genetic structure analyses using LD-pruned SNPs obtained through Plink (indep-pairwise parameters: window size 50, step 10, r^2 threshold 0.6). The Bayesian clustering analysis with Admixture v1.3.0 was performed to identify groups of genetically similar individuals, setting the potential number of distinct genetic groups from 1 to 9[61]. The optimal number of groups was determined by the lowest cross-validation (CV) error. To further confirm the genetic structure, principal component analysis (PCA) was conducted by using the GCTA software to evaluate the structure of the samples[62]. A phylogenetic tree was constructed to analyze the genetic relationships among individuals using the maximum likelihood (ML) algorithm in MEGA v12 (https://www.megasoftware.net/) software, employing the Kimura 2-parameter model with 1,000 bootstrap replications[63]. Given the strong pattern of isolation by distance (IBD), we further used the conStruct program to infer potential genetic clusters by considering the geographic distance of the samples[64]. We tested both non-spatial and spatial models with the number of clusters (K) varying from 1 to 10. Each experiment comprised four independent MCMC chains, each running for 5000 iterations. A cross-validation analysis with 8 replicates was conducted to compare the spatial and non-spatial models and determine the optimal K value.

Environmental association analysis

Nineteen bioclimatic variables with a spatial resolution of 30 arc seconds were acquired from the WorldClim database (http://www.worldclim.org/). To mitigate multicollinearity, environmental variables with high correlation (above 0.90) were excluded. The correlation analysis was performed using the R package "ade4". The remaining ten environmental variables used for subsequent analysis were: annual mean temperature (bio1), minimum temperature of the coldest month (bio6), temperature annual range (bio7), mean temperature of the wettest quarter (bio8), mean temperature of the warmest quarter (bio10), precipitation of the wettest month (bio13), precipitation of the driest month (bio14), precipitation seasonality (bio15), precipitation of the wettest quarter (bio16), and precipitation of the warmest quarter (bio18). To infer the relative contribution of environmental variables on genetic patterns, redundancy analysis (RDA) was conducted using the "rda" function in the R package "vegan"[65]. The SNP data were used as response variables, while the remaining ten environmental factors acted as explanatory variables. The Mantel test was conducted to assess the correlation between Fst/(1-Fst) and both geographic and environmental distances using the R v4.4.3 (https://www.r-project.org/) package "vegan", with significance determined through 999 permutations[65]. Geographic distances between sampling sites were calculated based on latitude and longitude using Euclidean distances. Environmental distances, incorporating ten bioclimatic variables, were also represented using Euclidean distances[66].

Environmentally adaptive genomic loci detection and annotation

Two strategies were utilized to detect variants associated with the environment. First, we used a univariate method for the latent factor mixed model (LFMM) to examine the correlation between genomic loci and environmental factors. The LFMM uses quick algorithms within a specific statistical model to analyze data while considering the influence of genetic structure[67]. The R v4.4.3 package "LEA" was used to conduct the analysis with the following parameters: 5 independent replicates, 10,000 iterations, 5,000 burn-in iterations, and latent factors set to K = 3 based on the genetic structure analysis. Loci with $|z| > 3$ and $P < 0.01$ were regarded as loci significantly associated with a particular environmental variable (EAL)[68]. We used BLAST v2.13.0 (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to align the EALs, including 150 bp upstream and downstream of each locus, with the genomes of *Calamus simplicifolius* (https://github.com/wy-cmd/Re-annotated-the-genomes-of-rattan), a species from the same genus as *C. rhabdocladus*, using an e-value of 1e-5[69]. To determine which specific biological processes the corresponding environment-associated candidate genes were enriched for, GO and KEGG (https://www.kegg.jp/kegg/mapper.html) enrichment analyses were conducted using OmicShare tools (http://www.omicshare.com/tools), based on Fisher's exact test and Q-value correction (adjusted P-value < 0.05). RDA can determine the amount of variance in response variables explained by multiple explanatory variables. This method was also used to identify genomic loci associated with local environmental adaptation. According to the method described by Sang et al.[70], the first four RDA axes, which explain 61% of the genetic variation, were selected to test the association between genomic loci and environmental factors. Loci with loadings in the distribution tails exceeding 3.5 standard deviations on the RDA axes were identified as significantly associated with environmental conditions. The RDA analysis was conducted using the R package "vegan"[65].

Data availability

All RAD-seq raw sequencing reads have been deposited at NCBI under the bio-project accession numbers PRJNA1129194. The SNP dataset generated and analyzed in this study has been submitted to Dryad and is currently under peer review (DOI: https://doi.org/10.5061/dryad.gf1vhhmz2). As the dataset has not yet been published, it is available from the corresponding author upon reasonable request (Email: liuguanglu@icbr.ac.cn).

Received: 13 November 2024; Accepted: 11 June 2025
Published online: 01 July 2025

References

1. Gadissa, F., Tesfaye, K., Dagne, K. & Geleta, M. Genetic diversity and population structure analyses of *Plectranthus Edulis* (Vatke) Agnew collections from diverse agroecologies in Ethiopia using newly developed EST-SSRs marker system. *BMC Genet.* **19**, 92–96. https://doi.org/10.1186/s12863-018-0682-z (2018).
2. Li, X. et al. De Novo transcriptome assembly and population genetic analyses for an endangered Chinese endemic *Acer miaotaiense*. *(Aceraceae) Genes*. **9**, 378. https://doi.org/10.3390/genes9080378 (2018).
3. Honnay, O. & Jacquemyn, H. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conserv. Biology*. **21**, 823–831. https://doi.org/10.1111/j.1523-1739.2006.00646.x (2007).
4. Ellis, J. R., Pashley, C. H., Burke, J. M. & McCauley, D. E. High genetic diversity in a rare and endangered sunflower as compared to a common congener. *Mol. Ecol.* **15**, 2345–2355. https://doi.org/10.1111/j.1365-294X.2006.02937.x (2006).
5. Toczydlowski, R. H. & Waller, D. M. Drift happens: molecular genetic diversity and differentiation among populations of jewelweed (*Impatiens capensis Merrb.*) reflect fragmentation of food plain forests. *Mol. Ecol.* **28**, 2495–2475. https://doi.org/10.1111/mec.15072 (2019).
6. Frankham, R., Bradshaw, C. J. A. & Brook, B. W. Genetics in conservation management: revised recommendations for the 50/500 rules, red list criteria and population viability analyses. *Conserv. Biology*. **170**, 56–63. https://doi.org/10.1016/j.biocon.2013.12.036 (2014).
7. Fan, S. H. et al. *Cultivation of Chinese Palm Rattan* (China Forestry Publishing House, 2022).
8. Jiang, Z. H. *World Bamboo and Rattan* (Liaoning Science and Technology, 2002).
9. Sarmah, P., Barua, P. K., Sarma, R. N., Sen, P. & Deka, P. C. Genetic diversity among Rattan genotypes from India based on RAPD-marker analysis. *Genet. Resour. Crop Evol.* **54**, 593–600. https://doi.org/10.1007/s10722-006-0017-1 (2007).
10. Li, Y. B. Study on assisted generation growth of rattan in Secondary Rain in Ganzhaling of Hainan, China. China: Doctoral's Thesis, Chinese Academy of Forestry (2019).
11. Nong, J. Q. Geographical differentiation of phenotypic traits of *Calamus rhabdocladus.* China: Master's Thesis, Chinese Academy of Forestry (2023).
12. Lin, Y. P., Lu, C. Y. & Lee, C. R. The Climatic association of population divergence and future extinction risk of *Solanum pimpinellifolium*. *AoB Plants*. **12**, plaa012. https://doi.org/10.1093/aobpla/plaa012 (2020).
13. Zuriaga, E. et al. Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genet. Resour. Crop Evol.* **56**, 39–51. https://doi.org/10.1007/s10722-008-9340-z (2009).
14. Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M. & Holderegger, R. A practical guide to environmental association analysis in landscape genomics. *Mol. Ecol.* **24**, 4348–4370. https://doi.org/10.1111/mec.13322 (2015).
15. Li, Z., Liang, F., Zhang, T. & P. & Enhanced tolerance to drought stress resulting from *Caragana korshinskii* CkWRKY33 in Transgenic Arabidopsis thaliana. *BMC Genom Data*. **22**, 11. https://doi.org/10.1186/s12863-021-00965-4 (2021).
16. Wang, Y. K., Hu, Y. & Zhang, T. Z. Current status and perspective of RAD-seq in genomic research. *Hereditas* **36**, 41–49. https://doi.org/10.3724/sp.j.1005.2014.00041 (2014).
17. Wang, S. et al. Sequence based ultra-dense genetic and physical maps reveal structural variations of allopolyploid cotton genomes. *Genome Biol.* **16**, 108–108. https://doi.org/10.1186/s13059-015-0678-1 (2015).
18. Liu, J. M. Landscape genomics reveal the molecular mechanisms of genetic differentiation in *Stipa breviflora* populations. China: Master's Thesis, Inner Mongolia University (2021).
19. Baird, N. A. et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*. **3**, e3376. https://doi.org/10.1371/journal.pone.0003376 (2008).
20. Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A. & Johnson, E. A. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* **17**, 240–248. https://doi.org/10.1101/gr.5681207 (2007).
21. Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A. & Johnson, E. A. Local de Novo assembly of RAD paired-end contigs using short sequencing reads. *PloS One*. **13**, e18561. https://doi.org/10.1371/journal.pone.0018561 (2011).
22. Gonen, S. et al. Linkage maps of the Atlantic salmon (*Salmo salar*) genome derived from RAD sequencing. *BMC Genom.* **15**, 166. https://doi.org/10.1186/1471-2164-15-166 (2014).
23. Brauner, S., Crawford, D. J. & Stuessy, T. F. Ribosomal DNA and RAPD variation in the rare plant family Lactoridaceae. *Amer J. Bot.* **79**, 1436–1439. https://doi.org/10.2307/2445144 (1992).
24. Gomory, D., Szczecińska, M., Sramko, G., Wołosz, K. & Sawicki, J. Genetic diversity and population structure of the rare and endangered plant species *Pulsatilla patens* (L.) mill in East central Europe. *PLoS ONE*. **11**, e0151730. https://doi.org/10.1371/journal.pone.0151730 (2016).
25. Yang, H. Study on genetic diversity and sex identification of *Daemonorops margaritae* & *Calamus simplificifolius*. China: Doctoral's Thesis, Chinese Academy of Forestry (2005).
26. Meena, R. K., Sharma, P., Bhandari, M. S. & Ginwal, H. S. Relatively higher genetic diversity with a moderate level of genetic differentiation was recorded for the sampled populations of *C. flagellum* distributed in North East Himalayas. *Indian Forester*. **145**, 535–543 (2019).
27. Sitisalwana, H., Wickneswari, R., Narwati, M. & Nur-supardi, M. N. Genetic diversity of naural populations of *Calamus palustris*. In:Proceeding:3rd National congress on genetics: Genetics towards better quality in life, Persyuan genetik. 18, 229–233 (1998).
28. Ramesha, B. T., Ravikanth, G., Rao, M. N., Ganeshaiah, K. N. & Umashaanker, R. Genetic structure of the Rattan *calamus thwaitesii* in core, buffer and peripheral regions of three protected areas in central Western ghats, india: do protected areas serve as refugia for genetic resources of economically important plants? *J. Genet.* **86**, 9–18. https://doi.org/10.1007/s12041-007-0002-2 (2007).
29. Priya, K., Indira, E., Sreekumar, V. B. & Renuka, C. Assessment of genetic diversity in *Calamus Vattayila* Renuka (Arecaceae) using ISSR markers. *J. Bamboo Ratten*. **15**, 61–69 (2016).
30. Tan, Z. Q., Leng, Y. & Zhou, H. T. RAPD analyses of 10 species in palmae. *J. Xiamen Univ. (Natural Science)*. **42**, 805–809 (2003).
31. Addisalem, A. B., Duminil, J., Wouters, D., Bongers, F. & Smulders, M. J. M. Fine-scale Spatial genetic structure in the frankincense tree *Boswellia papyrifera* (Del.) hochst. And implications for conservation. *Tree Genet. Genomes*. **12**, 86 (2016).
32. Hartvig, I. et al. Population genetic structure of the endemic Rosewoods *Dalbergia cochinchinensis* and *D. oliveri* at a regional scale reflects the Indochinese landscape and life-history traits. *Ecol. Evol.* **8**, 530–545. https://doi.org/10.1002/ece3.3626 (2018).
33. Zhai, S. H., Yin, G. S. & Yang, X. H. Population genetics of the endangered and wild edible plant *Ottelia acuminata* in Southwestern China using novel SSR markers. *Biochem. Genet.* **56**, 235–254 (2018).
34. Teixeira, T. M. & Nazareno, A. G. One step away from extinction: A population genomic analysis of A narrow endemic, tropical plant species. Front. *Plant. Sci.* **12**, 730258. https://doi.org/10.3389/fpls.2021.730258 (2021).
35. Ball, J. W. et al. Fine-scale species distribution modelling and genotyping by sequencing to examine hybridisation between two narrow endemic plant species. *Sci. Rep.* **10**, 1562. https://doi.org/10.1038/s41598-020-58525-2 (2020).
36. Ye, H. et al. Population genetic variation characterization of the boreal tree *Acer ginnala* in Northern China. *Sci. Rep.* **10**, 13515. https://doi.org/10.1038/s41598-020-70444-w (2020).
37. Lake, N. J. et al. Quantifying constraint in the human mitochondrial genome. *Nature* **635**, 391–400. https://doi.org/10.1038/s41586-024-08048-x (2024).
38. Lynch, M., Wei, W., Ye, Z. Q. & Pfrender, M. The genome-wide signature of short-term Temporal selection. *PNAS* **121**, e2307107121. https://doi.org/10.1073/pnas.2307107121 (2024).

39. Yang, J., Li, Y. & Miao, C. Y. Landscape genomics analysis of *Achyranthes bidentata* reveal adaptive genetic variations are driven by environmental variations relating to ecological habit. *Popul. Ecol.* **59**, 355–362. https://doi.org/10.1007/s10144-017-0599-9 (2017).
40. Hopken, M. W., Douglas, M. R. & Douglas, M. E. Stream hierarchy defines riverscape genetics of a North American desert fish. *Mol. Ecol.* **22**, 956–971. https://doi.org/10.1111/mec.12156 (2013).
41. Worth, J., Jordan, G. J., Marthick, J. R., McKinnon, G. E. & Vaillancourt, R. E. Chloroplast evidence for geographic stasis of the Australian bird-dispersed shrub *Tasmannia lanceolata* (Winteraceae). *Mol. Ecol.* **19**, 2949–2963. https://doi.org/10.1111/j.1365-294x.2010.04725.x (2010).
42. Tang, Y., Wu, W., Zheng, X. & H. & AT-Hook transcription factors show functions in *Liriodendron chinense* under drought stress and somatic embryogenesis. *Plants (Basel)*. **12**, 1353. https://doi.org/10.3390/plants12061353 (2003).
43. Rachel, G., Meirav, E. & Keren, M. Phylogeny and a structural model of plant MHX transporters. *BMC Plant Biol.* **12**, 133. https://doi.org/10.1186/1471-2229-13-75 (2013).
44. Annegret, K., Sandra, S. & Uwe, C. Benzothiadiazole-Induced priming for potentiated responses to pathogen infection, wounding, and infiltration of water into leaves requires the NPR1/NIM1 gene in Arabidopsis. *Plant Physiol.* https://doi.org/10.1104/pp.010744 (2002).
45. Fu, X., Tang, X. & Liu, W. Ubiquitination in plant biotic and abiotic stress. *Plant. Growth Regul.* **103**, 33–50. https://doi.org/10.1007/s10725-023-01095-w (2024).
46. Chen, Z. Palm plants. *Fujian Sci. Technol. Publishing House*. **9**, 77–78 (2001).
47. Editorial Committee of the Flora of China, Chinese Academy of Sciences. Flora of China. 13, 80–81. Beijing: Science Press. (1991).
48. Murray, M. G. & Thompson, W. F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **8**, 4321–4326. https://doi.org/10.1093/nar/8.19.4321 (1980).
49. Catchen, J., Hohenlohe, P. A. & Bassham, S. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140. https://doi.org/10.1111/mec.12354 (2013).
50. Guo, J. F., Wang, B. S. & Liu, Z. L. Low genetic diversity and population connectivity fuel vulnerability to climate change for the tertiary relict pine *Pinus bungeana*. *J. Syst. Evol.* **40**, 143–156. https://doi.org/10.1111/jse.12821 (2023).
51. Linck, E. & Battey, C. J. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol. Ecol. Resour.* **19**, 639–647. https://doi.org/10.1111/1755-0998.12995 (2019).
52. Lischer, H. E. L., Excoffier, L. & PGDSpider An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298–299. https://doi.org/10.1093/bioinformatics/btr642 (2012).
53. Nazareno, A. G., Dick, C. W. & Lohmann, L. G. Wide but not impermeable: testing the riverine barrier hypothesis for an Amazonian plant species. *Mol. Ecol.* **26**, 363–3648. https://doi.org/10.1111/mec.14142 (2017).
54. Nazareno, A. G., Dick, C. W. & Lohmann, L. G. A landscape genomic evaluation of wallace's riverine barrier hypothesis for three Amazon plant species. *Mol. Ecol.* **28**, 980–997. https://doi.org/10.1111/mec.14948 (2019).
55. Guzmán, S., Giudicelli, G. C. & Turchetto, C. Neutral and outlier single nucleotide polymorphisms disentangle the evolutionary history of a coastal Solanaceae species. *Mol. Ecol.* **31**, 2847–2864. https://doi.org/10.1111/mec.16441 (2022).
56. Nei, M. & Roychoudhury, A. K. Sampling variances of heterozygosity and genetic distance. *Genetics* **76**, 379–390 (1974).
57. Hudson, R. R., Boos, D. D. & Kaplan, N. L. A statistical test for detecting population subdivision. *Mol. Biol. Evol.* **9**, 138–151. https://doi.org/10.1093/oxfordjournals.molbev.a040703 (1992).
58. Wright, S. *Evolution and the Genetic of Population, Variability Within and among Natural Populations*vol. 4p.213–220 (University of Chicago Press, 1978).
59. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol. Ecol. Resour.* **10**, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x (2010).
60. Manni, F., Guerard, E. & Heyer, E. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using monmonier's algorithm. *Hum. Biol.* **76** https://doi.org/10.1353/hub.2004.0034 (2004). 173 – 90.
61. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry Estimation. *BMC Bioinform.* **12**, 1–6. https://doi.org/10.1186/1471-2105-12-246 (2011).
62. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
63. Li, M. Y. Population genetics and landscape genomics of the rare and endangered plant *Tetrastigma hemsleyanum* based on SLAF-seq technique. China: Master's Thesis, Henan Agricultural University (2021).
64. Bradburd, G. S. Inferring continuous and discrete population genetic structure across space. *Genetics* **210**, 33–52 (2018).
65. Oksanen, J., Blanchet, F. G., Friendly, M. & Kindt, R. Vegan: Community Ecology Package. R Package Version. (2017). Available at: https://CRAN.R-project.org/package=vegan vegan: Community Ecology Package. R package version 2.4.
66. Jiang, X. L., Gardner, E. M. & Meng, H. H. Land bridges in the pleistocene contributed to flora assembly on the continental Islands of South china: insights from the evolutionary history of *Quercus championii*. *Mol. Phylogenet. Evol.* **132**, 36–45. https://doi.org/10.1016/j.ympev.2018.11.021 (2019).
67. Frichot, E., Schoville, S. D., Bouchard, G. & François, O. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* **30**, 1687–1699. https://doi.org/10.1093/molbev/mst063 (2013).
68. Li, J. X. et al. Adaptive genetic differentiation in *Pterocarya stenoptera* (Juglandaceae) driven by multiple environmental variables were revealed by landscape genomics. *BMC Plant Biol.* **18**, 306. https://doi.org/10.1186/s12870-018-1524-x (2018).
69. Zhao, H. S. et al. The chromosome-level genome assemblies of two rattans (*Calamus simplicifolius* and *Daemonorops jenkinsiana*). *Gigascience* **7**, giy097. https://doi.org/10.1093/gigascience/giy097 (2018).
70. Sang, Y. P. et al. Genomic insights into local adaptation and future climate-induced vulnerability of a keystone forest tree in East Asia. *Nat. Commun.* **1**, 136541. https://doi.org/10.1038/s41467-022-34206-8 (2022).

Acknowledgements

We thank Junqing Nong for the sample collection. We also thank the anonymous reviewers and Associate Editors for their insightful comments. This study was supported by grants from the "14th Five-Year Plan" National Key R&D Program of China (No. 2021YFD2200501), International Bamboo and Rattan Center Fundamental Research Fund Special Project (No. 1632021021), and the International Bamboo and Rattan Center Sanya Research Base Key Discipline Construction Project (No. YJPY2024001-2).

Author contributions

R.G. conceptualized the study and developed the methodology. R.G. conducted the formal analysis. R.X. and S.W. curated the data. R.G. and R.X. wrote the original draft, while P.K.I., S.F., and G.L. contributed to review and editing. G.L. acquired funding. All authors reviewed and approved the final manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-06829-6.

Correspondence and requests for materials should be addressed to G.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.