

Article

Genetic Diversity Analysis of *Cymbidium eburneum* Lindl. (Orchidaceae) Based on SSR Markers

Feilong Hu ^{1,2,3,4}, Zhe Zhang ^{4,5} , Shunjiao Lu ^{1,2,3} , Zhiheng Chen ⁴, Haotian Zhong ^{1,2,3,4}, Liang Xi ^{1,2,3} and Guangsui Yang ^{1,2,3,*}

- ¹ Tropical Crop Germplasm Research Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China; h2210617716@163.com (F.H.); lushunjiao@catas.cn (S.L.); stiggg01@gmail.com (H.Z.); miai9275@gmail.com (L.X.)
- ² Key Laboratory of Crop Gene Resources and Germplasm Enhancement in Southern China, Ministry of Agriculture and Rural Affairs, Haikou 571101, China
- ³ Key Laboratory of Tropical Crops Germplasm Resources Genetic Improvement and Innovation of Hainan Province, The Engineering Technology Research Center of Tropical Ornamental Plant Germplasm Innovation and Utilization, Haikou 571101, China
- ⁴ Key Laboratory of Genetics and Germplasm Innovation of Tropical Special Forest Trees and Ornamental Plants, Ministry of Education, School of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, China; zhangzhe@hainanu.edu.cn (Z.Z.); 18982783311@163.com (Z.C.)
- ⁵ Sanya Institute of Breeding and Multiplication, Hainan University, Sanya 572025, China
- * Correspondence: ygungs@catas.cn

Abstract

Cymbidium eburneum Lindl. is a valuable ornamental orchid and breeding parent, but its genetic background remains unclear due to habitat destruction and germplasm mixing. This study developed specific SSR markers to evaluate the genetic diversity and structure of 96 *C. eburneum* Lindl. accessions from China and Vietnam. Transcriptome analysis identified 47,248 SSR loci. Sixteen polymorphic core primer pairs detected 150 alleles (mean $N_a = 9.375$) with an average Polymorphism Information Content (PIC) of 0.444. Observed heterozygosity ($H_o = 0.290$) was noticeably lower than expected ($H_e = 0.478$), indicating heterozygote deficiency. UPGMA clustering identified eight groups strongly correlated with geography. Principal Coordinate Analysis (PCoA) revealed a clear geographical differentiation pattern, featuring the most genetically cohesive group from Guangxi and more differentiated geographically marginal populations from Hainan and Vietnam. STRUCTURE analysis ($K = 2$) indicated two main gene pools with signals of genetic admixture. Geographical isolation was suggested as a potential driver of genetic differentiation. The Guangxi population represents a genetically consistent major reservoir, while marginal populations harbor unique variations. These findings provide a scientific basis for germplasm identification, conservation, and parental selection in *C. eburneum* Lindl. breeding.

Keywords: conservation genetics; population structure; endangered plant; germplasm resources; gene flow



Academic Editor: Kazuo Ichimura

Received: 24 March 2026

Revised: 13 April 2026

Accepted: 17 April 2026

Published: 21 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The Orchidaceae family represents one of the most evolutionarily distinct lineages in the plant kingdom, and the conservation and utilization of its germplasm resources remain focal points of botanical inquiry [1–4]. *C. eburneum* Lindl., a preeminent traditional Chinese orchid, possesses significant ornamental value. Phylogenetically, it belongs to the subgenus

Cyperorchis of the genus *Cymbidium* (Orchidaceae) and represents a basal taxon within this subgenus. Its native range encompasses Southwest China (Yunnan, Guangxi, Hainan) and the Indochina Peninsula (Vietnam, Myanmar, etc.), situated in the ecotone between tropical and subtropical climates [5]. Within the phylogeny of *Cymbidium*, *C. eburneum* Lindl. occupies a unique evolutionary position [6,7]. However, prolonged geographical isolation and habitat fragmentation have severely degraded wild resources, which may be reducing the genetic diversity and increasing the vulnerability of *C. eburneum* Lindl. populations, posing substantial challenges to conservation and breeding initiatives [8–11]. Consequently, a comprehensive assessment of the genetic diversity of *C. eburneum* Lindl., coupled with the establishment of a reliable molecular marker system, is critical for germplasm conservation, cultivar authentication, and molecular breeding.

Among various molecular marker technologies, orchid research necessitates techniques that function independently of whole-genome information, are operationally streamlined, and effectively delineate genetic discrepancies. Simple Sequence Repeat (SSR) markers align perfectly with these requisites [12]. SSR loci are derived from short tandem repeats within the genome or transcriptome; variation in repeat copy number facilitates the effective differentiation of genotypes. Accordingly, SSRs have emerged as one of the most prevalent marker technologies in plant molecular genetics. Furthermore, with advancements in high-throughput sequencing technologies, methodologies such as developing EST-SSRs from transcriptome data, mining genomic SSR loci, and evaluations of cross-species primer transferability have matured [13]. These approaches have been successfully applied to genetic investigations of various orchids, including *C. goeringi* and *C. faberi*, *C. ensifolium*, *Cattleya hybrids*, and *C. sinense* [14,15]. Compared to complex SNP/KASP arrays, EST-SSRs derived from transcriptomes are highly cost-effective, easily transferable across related taxa, and often linked to functional genes, making them highly suitable for evaluating non-model endangered species with limited genomic resources.

In the present study, we evaluated the genetic diversity and population structure of *C. eburneum* Lindl. utilizing SSR markers. Based on its fragmented distribution, we hypothesized that the long-term geographical barriers between mainland and island populations would drive strong genetic structuring, and prolonged habitat fragmentation might lead to an observable deficit of heterozygotes. Testing these hypotheses will provide a robust scientific basis for the germplasm conservation, cultivar authentication, and sustainable management of this endangered species.

2. Materials and Methods

2.1. Plant Materials and DNA Extraction

A total of 96 germplasm resources were collected, covering the major natural distribution areas and cultivation centers of *C. eburneum* Lindl. (Table 1). The collection included 81 accessions from the Guangxi population (Rongxian and Longzhou), representing the core distribution in mainland China; 3 accessions from the Hainan population (Wuzhishan, Exianling, and Ledong), representing island-isolated populations; 2 accessions from the Yunnan population (Dali, Gaoligongshan) and 3 from the Vietnam population (Northern and Southern regions), representing peripheral and transboundary populations. Additionally, 7 cultivars were included as references. Fresh young leaves were collected for all samples, rapidly dried with silica gel, and stored at -20°C [16]. DNA extraction was performed using a plant genomic DNA kit (Igebio, Wuhan, China), and DNA quality and concentration were verified using 1% agarose gel electrophoresis and a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), ensuring an A260/A280 ratio between 1.8 and 2.0 prior to PCR amplification. Given the endangered status and scattered distribution of wild *C. eburneum* Lindl. in marginal regions, our sampling scheme

is inevitably unbalanced, predominantly featuring the robust Guangxi population. Additionally, the seven cultivated varieties were included and analyzed together with the wild accessions in all population inferences (including UPGMA, PCoA, and STRUCTURE) to evaluate their genetic relationship and potential origin.

Table 1. Source distribution of the 96 *C. eburneum* Lindl. germplasm accessions used in this study.

Origin	No. of Samples
Guangxi	81
Hainan	3
Yunnan	2
Vietnam	3
Cultivated Variety	7
Total	96

2.2. SSR Primer Development, Screening, and PCR Amplification

Based on the Unigene sequences obtained from *C. eburneum* Lindl. transcriptome sequencing, MISA (Microsatellite identification tool) V1.0 software was used to search for SSR loci [17]. Screening criteria were set as follows: mono-nucleotide repeats ≥ 10 , di-nucleotide repeats ≥ 6 , and tri- to hexa-nucleotide repeats ≥ 5 . Based on the search results, Primer 3 v2.4.0 software was used to design and synthesize 96 pairs of specific primers [18]. DNA from eight typical germplasm accessions, selected to represent distinct geographical origins (Guangxi, Hainan, and Yunnan) and maximize genetic background coverage, was selected as templates for initial primer screening. PCR products were detected by 1% agarose gel electrophoresis. Finally, 16 pairs of core primers with clear amplification bands, rich polymorphism, and good reproducibility were selected for subsequent experiments. Primers were synthesized by Igebio (Wuhan) Biotechnology Co., Ltd. (Wuhan, China), with the 5' end of the forward primers labeled with FAM, HEX, TAMRA, or ROX fluorescent dyes. The total volume of the PCR reaction system was 20 μL , containing 10 μL of 2 $\times$ PCR MasterMix, 0.5 μL each of forward and reverse primers (10 $\mu\text{mol}\cdot\text{L}^{-1}$), 1 μL of template DNA (50 $\text{ng}\cdot\mu\text{L}^{-1}$), and 8 μL of ddH₂O. The PCR amplification program was: pre-denaturation at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 30 s; followed by a final extension at 72 °C for 10 min, and storage at 4 °C. Amplification products were detected using an ABI 3730xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), with GeneScan™ 500 LIZ™ Size Standard used as the internal standard.

2.3. Data Statistics and Analysis

GeneMarker V2.2.0 software was used to read the raw capillary electrophoresis data [19]. Allele sizes for each sample were determined based on internal standard peaks and manually verified. Since SSRs are codominant markers, the data were scored and coded as diploid genotypes (i.e., using a two-column format per locus recording the specific allele sizes) to meet the precise input requirements of GenAlEx and STRUCTURE software. Excel 2019 was used to calculate the number of alleles (N_a), effective number of alleles (N_e), Shannon's information index (I), observed heterozygosity (H_o), and expected heterozygosity (H_e) for each microsatellite locus. Polymorphism Information Content (PIC) was calculated using the formula $PIC = 1 - \sum P_i^2 - \sum \sum 2P_i^2P_j^2$ [20]. NTSYS-pc 2.10e software was used to calculate Nei's genetic distance and construct a phylogenetic tree using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [21]. Principal Coordinate Analysis (PCoA) was performed using GenAlEx 6.5 software [22,23]. Population genetic structure analysis was conducted using STRUCTURE 2.3.4 software, based

on the Admixture model and Correlated allele frequencies model [24,25]. The number of populations (K) was preset from 1 to 10, with 10 independent runs for each K value. The burn-in period was set to 100,000 iterations, followed by 100,000 MCMC iterations. The optimal number of populations was determined by calculating the ΔK value based on the Evanno method, implemented via the StructureSelector web server [26,27].

3. Results

3.1. Distribution Characteristics of SSR Loci in the *C. eburneum* Lindl Transcriptome

Simple Sequence Repeats (SSRs) were identified from the unigene sequences of the *C. eburneum* Lindl. transcriptome using MISA software. The search criteria were defined as a minimum of 10 repeats for mononucleotides, 6 for dinucleotides, and 5 for trinucleotides and higher. Adjacent SSRs separated by less than 100 bp were classified as compound SSRs. A total of 47,248 SSR loci were detected (Figure 1). The distribution was predominantly characterized by compound and mononucleotide repeats. Specifically, compound SSRs (types c and c*) were the most abundant, totaling 19,833 (41.98%), followed by mononucleotide repeats (p1) with 16,898 loci (35.76%). Dinucleotide (p2) and trinucleotide (p3) repeats accounted for 5401 (11.43%) and 4676 (9.90%), respectively, while tetra- to hexanucleotide motifs (p4–p6) were scarce, with a combined count of 440 (<1%). Overall, compound and mononucleotide repeats constituted more than three-quarters of the total identified loci, significantly outnumbering di-, tri-, and higher-order repeat motifs.

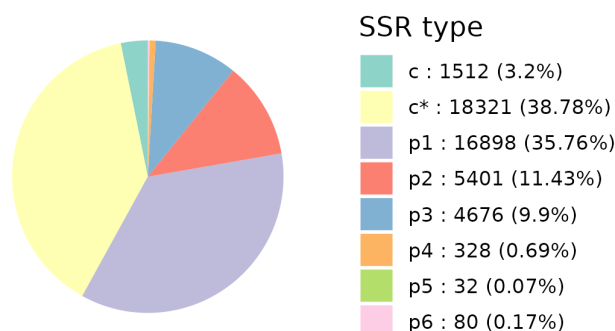


Figure 1. Distribution of SSR types in the *C. eburneum* Lindl. transcriptome.

3.2. Polymorphism Analysis

Based on the transcriptome SSR data, 96 primer pairs were designed and synthesized. Following PCR amplification and initial screening via agarose gel electrophoresis (Figure 2), 16 core primer pairs exhibiting clear amplification bands and high stability were selected to genotype 96 *C. eburneum* Lindl. germplasm accessions. Capillary electrophoresis data were analyzed using GeneMarker software (Figure 3), and genetic diversity parameters were calculated for each primer (Table 2). For certain loci (e.g., SSR108-6), a few samples failed to amplify consistently after repeated attempts; these were treated as missing data in subsequent statistical analyses.

The results showed that a total of 150 alleles were amplified by the 16 SSR markers. The number of alleles (N_a) per locus ranged from 3 (SSR108-20) to 16 (SSR108-1, SSR108-26), with an average of 9.375 alleles per locus. The effective number of alleles (N_e) varied from 1.136 to 3.026, with an average of 2.096; locus SSR108-43 exhibited the highest N_e value. Regarding genetic diversity parameters, Shannon's Information Index (I) ranged from 0.311 to 1.555, with an average of 1.041. The mean observed heterozygosity (H_o) and expected heterozygosity (H_e) were 0.290 and 0.478, respectively.

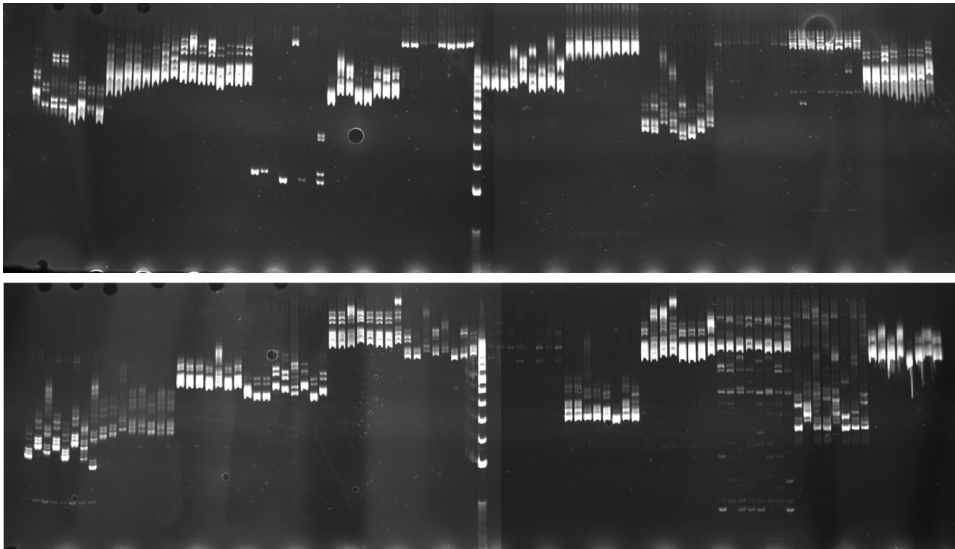


Figure 2. Electrophoresis band patterns of selected primers.

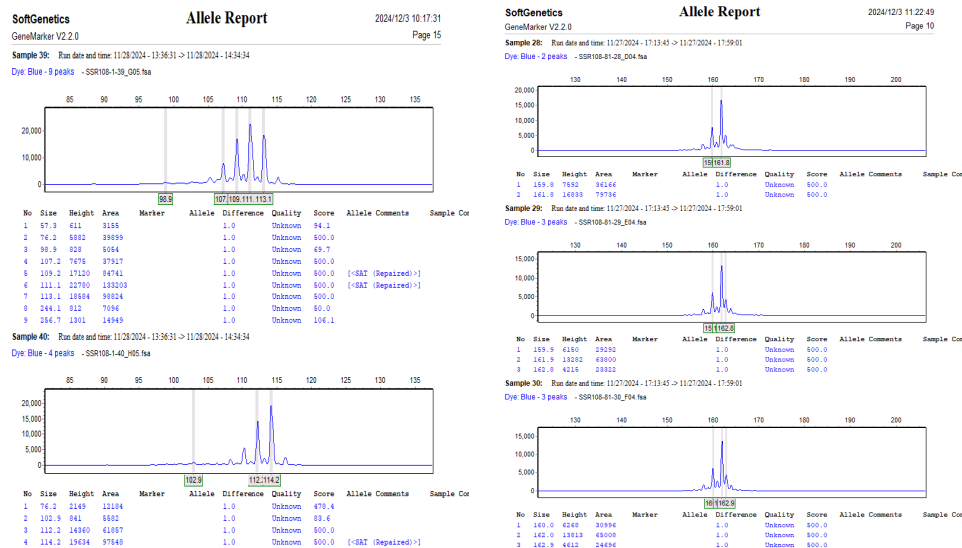


Figure 3. Capillary electrophoresis profiles of selected samples amplified by SSR primers.

Table 2. Genetic diversity parameters of 16 SSR primer pairs.

Primer ID	Number of Samples	<i>Na</i>	<i>Ne</i>	<i>I</i>	<i>Ho</i>	<i>He</i>	<i>PIC</i>
SSR108-1	96	16	2.573	1.538	0.240	0.611	0.594
SSR108-4	96	7	1.379	0.677	0.135	0.275	0.268
SSR108-6	88	15	2.933	1.451	0.852	0.659	0.603
SSR108-11	96	12	2.889	1.457	0.729	0.654	0.621
SSR108-13	96	12	2.562	1.347	0.479	0.610	0.568
SSR108-20	96	3	1.811	0.663	0.01	0.448	0.352
SSR108-23	96	8	1.827	0.995	0.25	0.453	0.428
SSR108-26	96	16	2.519	1.555	0.177	0.603	0.588
SSR108-43	96	5	3.026	1.197	0.302	0.669	0.604
SSR108-47	96	4	1.500	0.683	0.281	0.333	0.314
SSR108-55	96	7	1.980	0.954	0.177	0.495	0.441
SSR108-64	96	8	1.924	1.036	0.229	0.480	0.455
SSR108-77	96	10	2.395	1.123	0.375	0.582	0.511

Table 2. Cont.

Primer ID	Number of Samples	<i>Na</i>	<i>Ne</i>	<i>I</i>	<i>Ho</i>	<i>He</i>	<i>PIC</i>
SSR108-81	96	13	1.831	1.133	0.292	0.454	0.439
SSR108-93	96	6	1.136	0.311	0.031	0.120	0.117
SSR108-95	96	8	1.256	0.530	0.083	0.204	0.200
Mean	95.5	9.375	2.096	1.041	0.290	0.478	0.444

Note: *Na*, number of alleles; *Ne*, effective number of alleles; *I*, Shannon’s information index; *Ho*, observed heterozygosity; *He*, expected heterozygosity; *PIC*, polymorphism information content.

Polymorphism Information Content (*PIC*) is a critical indicator for evaluating the degree of variation in SSR markers. In this study, *PIC* values for the 16 loci ranged from 0.117 to 0.621, with an average of 0.444 (Table 2). Specifically, seven loci—SSR108-1, SSR108-6, SSR108-11, SSR108-13, SSR108-26, SSR108-43, and SSR108-77—had *PIC* values greater than 0.500, indicating high polymorphism. With the exception of SSR108-93 and SSR108-95, the remaining loci were classified as moderately polymorphic ($0.25 < PIC < 0.5$).

3.3. Phylogenetic Analysis of *C. eburneum* Lindl. Germplasm Resources

To systematically elucidate the phylogenetic relationships among the 96 *C. eburneum* Lindl. germplasm accessions, a phylogenetic tree was constructed using the UPGMA method based on genetic distances derived from SSR markers (Figure 4). The results indicated distinct genetic differentiation within the population, which clustered topologically into eight major groups (Cluster 0–Cluster 7).

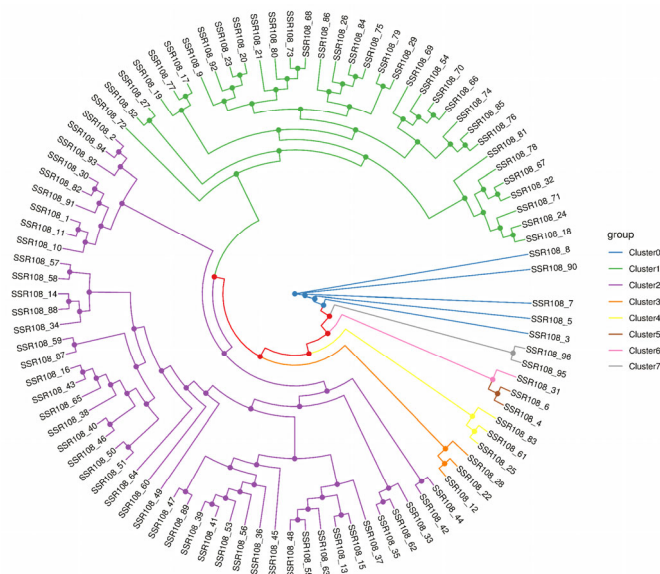


Figure 4. UPGMA clustering tree of 96 *C. eburneum* Lindl. accessions based on 16 SSR markers.

Cluster 2 emerged as the largest core group on one side of the phylogenetic tree, comprising 46 accessions. This cluster was characterized by tight internal nodes and short branch lengths, indicating generally small genetic distances and high genetic similarity among its members. The accessions in this clade predominantly originated from Rongxian, Guangxi. This was followed by the closely related Cluster 1, which contained 31 accessions. As shown in the phylogenetic tree, Cluster 1 and Cluster 2 are grouped together into a major clade, suggesting a close genetic relationship between them. However, Cluster 1 formed a relatively independent sub-clade in the topology, indicating that while sharing a partial genetic background with Cluster 2, it has undergone a certain degree of genetic differentiation driven by distinct ecological environments or artificial selection pressures.

In addition, Clusters 3–7, located at the peripheral or basal positions of the phylogenetic tree, constituted several distinct groups. These clades were primarily composed of germplasm from Hainan, Yunnan, and Vietnam. They exhibited longer branch lengths and greater genetic distances from the most genetically cohesive group of Guangxi Rongxian. Notably, the UPGMA dendrogram revealed that Cluster 0 formed a highly distinct basal clade, separating early from the main wild populations. The cultivated varieties were primarily positioned within this independent cluster, indicating that they possess a unique genetic background, likely driven by long-term artificial selection and cultivation practices.

3.4. Spatial Genetic Structure Based on Principal Coordinate Analysis (PCoA)

To further validate the clustering results and elucidate the genetic structure of *C. eburneum* Lindl. germplasm, from a spatial perspective, Principal Coordinate Analysis (PCoA) was performed based on Nei's genetic distance matrix. The results showed (Figure 5) that the first two principal coordinates, PCoA 1 and PCoA 2, explained 43.45% and 13.39% of the total genetic variation, respectively. The cumulative variance explained reached 56.84%, effectively capturing the major genetic relationships among the populations.

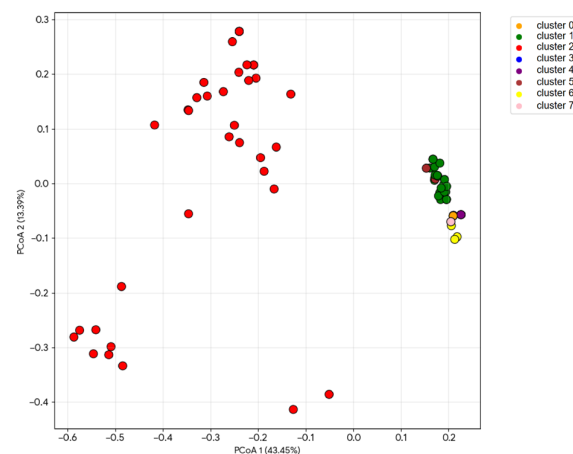


Figure 5. Principal Coordinate Analysis (PCoA) scatter plot of 96 *C. eburneum* Lindl. germplasm accessions based on 16 SSR markers.

The two-dimensional PCoA scatter plot revealed a distinct separation along the first principal coordinate (PCoA 1). The most genetically cohesive group, predominantly composed of Cluster 2, occupied the negative region on the left side of the plot. Although Cluster 2 contained the largest number of accessions and was grouped into a single major clade in the phylogenetic tree, it exhibited a wide and diffuse distribution in the PCoA plot. This high degree of dispersion suggests that the Rongxian (Guangxi) population has accumulated substantial internal genetic differentiation over a long reproductive history. This pattern implies that Rongxian may serve as a potential diversity hotspot for *C. eburneum* Lindl., reflecting a large effective population size and a prolonged evolutionary history.

In contrast, Cluster 1 and Clusters 3–7 were primarily distributed in the positive region on the right side, forming a tightly clustered genetic group. This distribution pattern highlights their evolutionary relationship with the core group. Although the accessions on the right side originated from diverse geographic locations, their collective separation from the core group along the horizontal axis suggests that they may have derived from a secondary center of origin that differentiated from the most genetically cohesive group.

3.5. Population Genetic Structure Analysis

The genetic structure of 96 *C. eburneum* Lindl. germplasm accessions was analyzed using the Bayesian model implemented in the STRUCTURE software. The ΔK statistic (Figure 6) revealed a distinct peak at $K = 2$, indicating that the tested population is optimally divided into two gene pools.

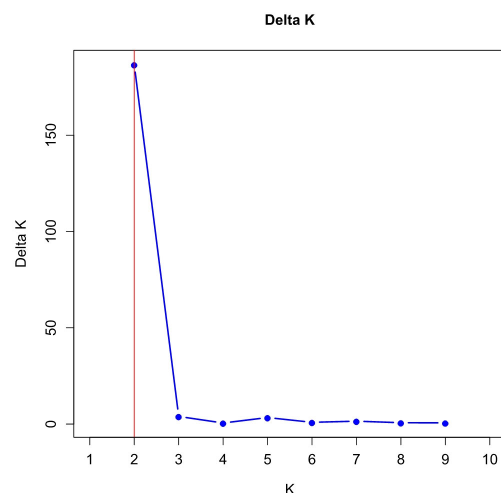


Figure 6. Estimation of the optimal number of clusters K using the ΔK statistic.

At $K = 2$ (Figure 7), the population exhibited significant differentiation. The dominant “blue” gene pool encompassed the vast majority of individuals in Clusters 1 and 2 (originating from Rongxian, Guangxi), indicating a highly consistent genetic background and a high level of homozygosity within the germplasm of this region. In contrast, the “red” gene pool was primarily distributed among the distinct accessions in Clusters 3–7, which originated from Hainan, Yunnan, and Vietnam, reflecting significant genetic divergence. Additionally, certain individuals (e.g., Ce13) displayed an admixed genetic constitution (a mosaic of red and blue components). This admixture likely results from gene exchange or hybridization between geographically distinct populations, providing evidence of gene flow during the evolutionary history of *C. eburneum* Lindl. germplasm resources.

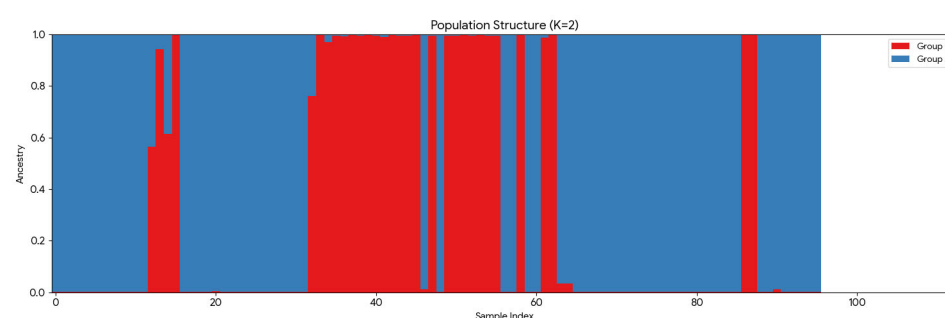


Figure 7. Bayesian clustering of 96 *C. eburneum* Lindl. germplasm accessions at $K = 2$.

4. Discussion

4.1. Effectiveness and Advantages of SSR Markers in *C. eburneum* Lindl.

Molecular markers are powerful tools for analyzing the genetic background of germplasm resources. The SSR markers developed in this study, based on the *C. eburneum* Lindl. transcriptome, demonstrated good stability and applicability. The 16 selected core primer pairs achieved a 100% amplification success rate and polymorphic locus rate across 96 *C. eburneum* Lindl. DNA samples, with an average PIC value of 0.444. This value is comparable to or slightly higher than the PIC value (0.345) reported by Shen et al. (2025) using

KASP markers in *C. sifolium*, and similar to the result ($PIC = 0.48$) reported by Ai et al. (2019) in *Cymbidium ensifolium* SSR research [28,29]. This may be because the SSR markers in this study were derived from transcriptome coding regions, where sequences are relatively conserved due to functional constraints [30,31]. Nevertheless, the PIC values of loci such as SSR108-11 and SSR108-6 exceeded 0.6, and the average number of alleles ($N_a = 9.375$) was at a relatively high level, sufficient to meet the needs of fine-scale genetic variation detection within *C. eburneum* Lindl. This also provides potential universal marker resources for comparative genomics research in the genus *Cymbidium*. These results demonstrate that the selected primers possess high polymorphism detection capability and can effectively reveal the level of genetic variation within *C. eburneum* Lindl. resources.

4.2. Genetic Diversity Level of *C. eburneum* Lindl. Germplasm Resources

Genetic diversity is the foundation for species adaptation and evolution. This study revealed that the expected heterozygosity (H_e) of the *C. eburneum* Lindl. population was 0.478, and Shannon's information index (I) was 1.041. Compared with other *C.* species, this diversity level is higher than that reported by Liang et al. (2021) for *C. faberi* in the Qinling Mountains ($H_e = 0.396$), and comparable to the results reported by Li et al. (2023) for *C. sinense* and its kindred germplasms ($H_e = 0.597$), indicating an overall moderately high level [32,33]. This may be related to the fact that *C. eburneum* Lindl. is a typical insect-pollinated plant that tends to outcross under natural conditions, thereby maintaining high genetic variation [34–36]. Notably, the observed heterozygosity ($H_o = 0.290$) in this study was significantly lower than the expected heterozygosity ($H_e = 0.478$), indicating a clear heterozygote deficiency within the population. The observation that H_o was significantly lower than H_e indicates an excess of homozygotes in the population. This result is similar to findings by Hu et al. (2024) in the endangered plant *Saussurea*, which may be related to inbreeding depression or small population effects [37]. Furthermore, this heterozygote deficit could also be attributed to the Wahlund effect (caused by pooling geographically distinct subpopulations with different allele frequencies into a single analysis) or the potential presence of null alleles across the analyzed loci.

4.3. Genetic Structure Differentiation and Conservation Strategies

Population structure analysis is key to formulating germplasm conservation strategies. Through UPGMA clustering, PCoA, and STRUCTURE analysis ($K = 2$), this study collectively revealed a significant “Core-Peripheral” dual differentiation pattern in *C. eburneum* Lindl. germplasm resources. Figure 7 clearly shows that the blue gene pool, represented by the Rongxian (Guangxi) population, occupies the core position, while the “red gene pool,” represented by populations from Hainan, Yunnan, and Vietnam, is distributed in marginal specific areas. This high degree of genetic differentiation is strongly correlated with geographical origin, providing observational support for the hypothesis that “geographical isolation is the main driver of genetic differentiation” [38–41]. The geographical barrier of the Qiongzhou Strait and the complex terrain of the Indochina Peninsula may have blocked gene flow between different geographical populations, leading to independent genetic evolution. Despite the geographic and artificial barriers that separate the cultivated varieties (such as those in Cluster 0) from the wild groups, our analysis indicated that these cultivars share the closest genetic affinity with the wild populations from Hainan. This suggests that the Hainan germplasm might have served as a foundational source for early domestication and breeding of these cultivated varieties.

Author Contributions: Conceptualization, F.H. and G.Y.; methodology, F.H.; validation, F.H. and S.L.; investigation, F.H., S.L., Z.C., H.Z. and L.X.; resources, Z.Z. and G.Y.; data curation, F.H.; writing—original draft preparation, F.H.; writing—review and editing, Z.Z. and G.Y.; supervision, Z.Z. and G.Y.; project administration, Z.Z. and G.Y.; funding acquisition, G.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2021YFD1200205.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Freudenstein, J.V. Orchid Phylogenetics and Evolution: History, Current Status and Prospects. *Ann. Bot.* **2025**, *135*, 805–822. [[CrossRef](#)] [[PubMed](#)]
- Fay, M.F. Orchid Conservation: How Can We Meet the Challenges in the Twenty-First Century? *Bot. Stud.* **2018**, *59*, 16. [[CrossRef](#)] [[PubMed](#)]
- Liu, Q.; Chen, J.; Corlett, R.T.; Fan, X.; Yu, D.; Yang, H.; Gao, J. Orchid Conservation in the Biodiversity Hotspot of Southwestern China. *Conserv. Biol.* **2015**, *29*, 1563–1572. [[CrossRef](#)] [[PubMed](#)]
- Wraith, J.; Pickering, C. Quantifying Anthropogenic Threats to Orchids Using the IUCN Red List. *Ambio* **2018**, *47*, 307–317. [[CrossRef](#)]
- Zhang, G.-Q.; Chen, G.-Z.; Chen, L.-J.; Zhai, J.-W.; Huang, J.; Wu, X.-Y.; Li, M.-H.; Peng, D.-H.; Rao, W.-H.; Liu, Z.-J.; et al. Phylogenetic Incongruence in Cymbidium Orchids. *Plant Divers.* **2021**, *43*, 452–461. [[CrossRef](#)]
- Chen, Z.; Gao, L.; Wang, H.; Feng, S. Molecular Identification and Phylogenetic Analysis of *Cymbidium* Species (Orchidaceae) Based on the Potential DNA Barcodes *matK*, *rbcl*, *psbA-trnH*, and Internal Transcribed Spacer. *Agronomy* **2024**, *14*, 933. [[CrossRef](#)]
- Wang, M.; Chen, J.-B.; Chen, G.-Z.; Chen, L.-J.; Wu, X.-Y.; Huang, J. The Complete Chloroplast Genome of *Cymbidium eburneum* (Orchidaceae). *Mitochondrial DNA Part B* **2019**, *4*, 2359–2360. [[CrossRef](#)]
- Sharma, S.K.; Kumaria, S.; Tandon, P.; Rao, S.R. Single Primer Amplification Reaction (SPAR) Reveals Inter- and Intra-Specific Natural Genetic Variation in Five Species of *Cymbidium* (Orchidaceae). *Gene* **2011**, *483*, 54–62. [[CrossRef](#)]
- Rajeswara Rao, B.R. Genetic Diversity, Genetic Erosion, Conservation of Genetic Resources, and Cultivation of Medicinal Plants. In *Genetic Diversity and Erosion in Plants*; Ahuja, M.R., Jain, S.M., Eds.; Sustainable Development and Biodiversity; Springer International Publishing: Cham, Switzerland, 2016; Volume 8, pp. 357–407.
- Zhou, Z.; Shi, R.; Zhang, Y.; Xing, X.; Jin, X. Orchid Conservation in China from 2000 to 2020: Achievements and Perspectives. *Plant Divers.* **2021**, *43*, 343–349. [[CrossRef](#)]
- Sarmah, D.; Gurung, N.; Saini, P.; Devi, M.P.; Mohapatra, P.; Pramanik, K.; Jena, C.; B., R. Advances in Orchid Conservation: Biotechnology Applications and Global Trade Dynamics. *Front. Conserv. Sci.* **2025**, *6*, 1653827. [[CrossRef](#)]
- Cheng, M.; Liu, Y.; Jin, C.; Jiang, X.; Chen, X.; Wang, F.; Duan, Y.; Zhuo, X.; Peng, D. Genome-Wide Identification of SNP and SSR Markers from *Cymbidium goeringii* and *C. faberi* for Their Potential Application in Breeding. *Horticulturae* **2025**, *11*, 622. [[CrossRef](#)]
- Ouyang, P.; Kang, D.; Mo, X.; Tian, E.; Hu, Y.; Huang, R. Development and Characterization of High-Throughput EST-Based SSR Markers for *Pogostemon cablin* Using Transcriptome Sequencing. *Molecules* **2018**, *23*, 1014. [[CrossRef](#)] [[PubMed](#)]
- Kim, K.C.; Kang, S.; Kim, S.-L.; Perera, R.M.T.D.; Woo, J.K.; Sandunika, K.H.; Kim, J.-H.; Lee, D.-S. Development of Simple Sequence Repeat (SSR) Markers from a Genome Survey of a *Cymbidium kanran* Makino Population in Jeju Island, Republic of Korea. *J. Microbiol. Biotechnol.* **2025**, *35*, e2501013. [[CrossRef](#)] [[PubMed](#)]
- Tsai, C.-C.; Shih, H.-C.; Wang, H.-V.; Lin, Y.-S.; Chang, C.-H.; Chiang, Y.-C.; Chou, C.-H. RNA-Seq SSRs of Moth Orchid and Screening for Molecular Markers across Genus *Phalaenopsis* (Orchidaceae). *PLoS ONE* **2015**, *10*, e0141761. [[CrossRef](#)]
- Fernández, J.F.; Sork, V.L.; Gallego, G.; López, J.; Bohorques, A.; Tohme, J. Cross-Amplification of Microsatellite Loci in a neotropical *Quercus* Species and Standardization of DNA Extraction from Mature Leaves Dried in Silica Gel. *Plant Mol. Biol. Report.* **2000**, *18*, 397. [[CrossRef](#)]
- Thiel, T.; Michalek, W.; Varshney, R.; Graner, A. Exploiting EST Databases for the Development and Characterization of Gene-Derived SSR-Markers in Barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)]
- Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—New Capabilities and Interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [[CrossRef](#)]
- Hulce, D.; Li, X.; Snyder-Leiby, T.; Johathan Liu, C.S. GeneMarker® Genotyping Software: Tools to Increase the Statistical Power of DNA Fragment Analysis. *J. Biomol. Tech. JBT* **2011**, *22*, S35–S36.

20. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am. J. Hum. Genet.* **1980**, *32*, 314–331.
21. Rohlf, F.J. *NTSYS-Pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1: Owner Manual*; Exeter Software: New York, NY, USA, 2000.
22. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—An Update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)]
23. Peakall, R.; Smouse, P.E. GenAlEx 6: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research. *Mol. Ecol. Notes* **2006**, *6*, 288–295. [[CrossRef](#)]
24. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **2000**, *155*, 945–959. [[CrossRef](#)] [[PubMed](#)]
25. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics* **2003**, *164*, 1567–1587. [[CrossRef](#)] [[PubMed](#)]
26. Li, Y.L.; Liu, J.X. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* **2018**, *18*, 176–177. [[CrossRef](#)]
27. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)]
28. Ai, Y.; Chen, L.; Xie, T.X.; Chen, J.; Lan, S.R.; Peng, D.H. Construction of Core Collection of *Cymbidium ensifolium* Cultivars Based on SSR Fluorescent Markers. *Acta Hortic. Sin.* **2019**, *46*, 1999–2008. [[CrossRef](#)]
29. Shen, B.; Shen, A.; Tan, Y.; Liu, L.; Li, S.; Tan, Z. Development of KASP Markers, SNP Fingerprinting and Population Genetic Analysis of *Cymbidium ensifolium* (L.) Sw. Germplasm Resources in China. *Front. Plant Sci.* **2025**, *15*, 1460603. [[CrossRef](#)]
30. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic Microsatellite Markers in Plants: Features and Applications. *Trends Biotechnol.* **2005**, *23*, 48–55. [[CrossRef](#)]
31. Eujayl, I.; Sorrells, M.E.; Baum, M.; Wolters, P.; Powell, W. Isolation of EST-Derived Microsatellite Markers for Genotyping the A and B Genomes of Wheat. *Theor. Appl. Genet.* **2002**, *104*, 399–407. [[CrossRef](#)]
32. Li, M.; Luo, Q.; Zhao, X.; Ding, C.; Chi, Z.; Lu, Z.; Bo, W. Divergence and Relationships of *Cymbidium Tortisepalum* and Its Kindred Germplasms by EST-SSR Markers. *J. Hortic. Sci. Biotechnol.* **2023**, *98*, 483–494. [[CrossRef](#)]
33. Liang, H.-Y.; Wang, X.-G.; Chen, W.; Hu, S.-N.; Yang, F.; Yang, X.-T. Spatial Genetic Structure of Terrestrial Orchid *Cymbidium faberi* in the Qinling Mountains Revealed by Microsatellite Loci. *Plant Syst. Evol.* **2021**, *307*, 5. [[CrossRef](#)]
34. Hamrick, J.L.; Godt, M.J.W. Effects of Life History Traits on Genetic Diversity in Plant Species. *Philos. Trans. Biol. Sci.* **1996**, *351*, 1291–1298. [[CrossRef](#)]
35. Dharmarajan, G.; Beatty, W.S.; Rhodes, O.E. Heterozygote Deficiencies Caused by a Wahlund Effect: Dispelling Unfounded Expectations. *J. Wildl. Manag.* **2013**, *77*, 226–234. [[CrossRef](#)]
36. Suetsugu, K. Autonomous Self-Pollination and Insect Visitors in Partially and Fully Mycoheterotrophic Species of *Cymbidium* (Orchidaceae). *J. Plant Res.* **2015**, *128*, 115–125. [[CrossRef](#)] [[PubMed](#)]
37. Hu, L.; Wang, J.; Wang, X.; Zhang, D.; Sun, Y.; Lu, T.; Shi, W. Development of SSR Markers and Evaluation of Genetic Diversity of Endangered Plant *Saussurea involucreata*. *Biomolecules* **2024**, *14*, 1010. [[CrossRef](#)] [[PubMed](#)]
38. Verbylaitė, R.; Aravanopoulos, F.A.; Baliuckas, V.; Tourvas, N.; Farsakoglou, A.-M.; Kotina, V.-M.; Lyrou, F.G.; Juškauskaitė, A.; Petrokas, R.; Lygis, V. The Central-Periphery Hypothesis Revisited: Implications for Long-Term Genetic Conservation. *Plants* **2025**, *14*, 3563. [[CrossRef](#)]
39. Csilléry, K.; Yang, H.; Hung, T.H.; Rodríguez-Rodríguez, P.; Miller, J.; Mantovani, V.; Chauvier-Mendes, Y.; Nobis, M.P. GenDivRange: A Global Dataset of Geo-Referenced Population Genetic Diversity across Species Ranges. *Sci. Data* **2025**, *12*, 980. [[CrossRef](#)]
40. Eckert, C.G.; Samis, K.E.; Loughheed, S.C. Genetic Variation across Species' Geographical Ranges: The Central–Marginal Hypothesis and Beyond. *Mol. Ecol.* **2008**, *17*, 1170–1188. [[CrossRef](#)]
41. Slatkin, M. Gene Flow and the Geographic Structure of Natural Populations. *Science* **1987**, *236*, 787–792. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.