horticulturae

MDPI

Article

ddRAD-seq Reveals Genetic Diversity and Population Structure of *Primula beesiana*

Qishao Li [1,2], Zihan Li [1,2], Sihan Zhang [1,2], Zhanghong Dong [1,2], Yongpeng Ma [3,*] and Peiyao Xin [1,2,*]

[1] Yunnan Key Laboratory of Landscape Plant Resource Cultivation and Application, Kunming 650224, China; lqs991104@swfu.edu.cn (Q.L.); lizihan966@163.com (Z.L.); 15752084415@163.com (S.Z.); dongzhanghong@swfu.edu.cn (Z.D.)

[2] Yunnan Province Engineering Research Center for Functional Flower Resources and Industrialization, Kunming 650224, China

[3] Yunnan Provincial Key Laboratory for Integrated Conservation of Extremely Small Populations of Wild Plants, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

* Correspondence: mayongpeng@mail.kib.ac.cn (Y.M.); xpytgyx@163.com (P.X.)

Abstract

Primula beesiana is a perennial herbaceous plant predominantly distributed in the alpine wetland regions of Yunnan Province, China. This species faces dual threats from habitat fragmentation and climate change, but research into its genetic background is severely lacking. Consequently, systematic analysis of the genetic diversity and population structure of *Primula beesiana* is crucial in formulating scientific conservation strategies. In this study, 86 individuals from six natural populations in Lijiang City, Yunnan Province, were collected and genotyped using double-digest restriction site-associated DNA sequencing (ddRAD-seq). A total of 1537 high-quality SNP loci were identified and used for genetic diversity, principal component (PCA), population structure (STRUCTURE), and gene flow analyses. Analysis of base substitutions revealed twelve mutation types, with transversions accounting for 67.9% and a transition/transversion ratio (Ti/Tv) of 0.47, potentially indicating strong environmental selection pressure. Although high overall genetic diversity was observed, significant genetic differentiation may exist among populations (Fst = 0.0056-0.0407), with heterozygote deficiency detected across all populations. Genetic structure analyses consistently grouped the six populations into four distinct clusters. Populations MDJ, WH, and HS each formed independent clusters, exhibiting clear genetic isolation, whereas XHC2, XHC1, and NX clustered together, showing high genetic similarity and frequent gene flow. Mantel tests demonstrated a significant positive correlation between genetic and geographical distances ($r = 0.854$, $p < 0.01$), supporting an isolation-by-distance model. Gene flow estimates varied considerably among populations (5.90-44.69) and decreased with increasing geographical distance. This study provides the first genomic-level evidence of significant genetic differentiation and isolation based on distance in *Primula beesiana* populations, offering crucial scientific support in identifying evolutionarily significant units and developing zoned conservation management strategies for this species.

Keywords: *Primula beesiana* Forrest; genetic diversity; population structure; ddRAD-seq; single-nucleotide polymorphism (SNP)

Check for updates

Academic Editor: Luigi De Bellis

Received: 27 November 2025
Revised: 27 January 2026
Accepted: 30 January 2026
Published: 31 January 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) license.

1. Introduction

Primula beesiana is a perennial herbaceous plant characterized by a flowering period concentrated between June and July [1]. Recognized for its vividly colored and abundant blooms, this species has considerable ornamental value and breeding potential. Its natural distribution is primarily centered in the high-altitude regions of northern and northwestern Yunnan Province, China, where it is typically found in wet grasslands, streamside areas, marshy meadows, and open grasslands at elevations of 2400–3200 m [1]. In terms of cultivation, it is often grown as a potted plant in greenhouses during early spring in northern China, while in southern regions, it can be cultivated outdoors in flowerbeds or on semi-shaded slopes. Notably, *Primula beesiana* demonstrates superior tolerance to high-temperature stress compared to its relative, *Primula secundiflora*. *Primula beesiana* holds significant horticultural value, with current commercial use centered on ornamental and landscape development [2]. Its distinctive umbellate inflorescence and multi-layered whorls of blush-pink flowers (Figure 1) create a striking visual display, making it a favored choice for flower beds, streamside planting, and specialty garden design. As a summer-blooming primrose endemic to high-altitude regions of southwestern China, it can effectively enrich floral diversity in summer gardens in cool climates and allow ornamental plant displays to attract interest for a longer, cross-seasonal period.



Figure 1. Morphological characteristics of *Primula beesiana*.

This species shows considerable ecological adaptability, performing well in cool, moist environments. Under appropriate horticultural conditions, it is relatively easy to cultivate and maintain, and it can be propagated via division or seed, showing potential for broader landscape use. Moreover, its vivid flower color and upright inflorescence structure represent valuable parental traits for hybridization within the genus *Primula*, providing important genetic resources for breeding novel cultivars with enhanced ornamental appeal and wider adaptability.

Primula beesiana, despite its significant ecological and economic value, continues to face substantial survival pressures in its native habitats due to combined threats from human activities and climate change [3]. These habitats are situated at the junction of the Qinghai–Tibet Plateau and the Hengduan Mountains, a region characterized by a complex topography, pronounced altitudinal variation, and diverse climatic conditions. Such environmental heterogeneity supports rich ecosystems and high biodiversity [4]. Within this landscape, alpine wetlands and streamside habitats offer relatively suitable growing conditions for *Primula beesiana*. However, the region's distinctive geoclimatic features also contribute to highly fragmented and spatially constrained habitats, increasing susceptibility to environmental fluctuations. These factors collectively impair the species' adaptive capacity at the genetic level [5]. In recent years, intensified human disturbances—such as overgrazing, tourism infrastructure development, and the effects of climate change—have

further reduced inter-population gene flow [6]. This decline may exacerbate genetic drift and alter spatial patterns of genetic diversity. As a result, habitat fragmentation has become more severe, accompanied by gradual declines in population sizes and distribution ranges [3]. Moreover, *Primula beesiana* is not included in China's Catalogue of Key Protected Wild Plants, and no systematic conservation strategy has been established for its preservation. With it being a morphologically distinctive species within the genus *Primula*, previous studies have largely focused on morphological description and cultivation techniques. To date, there is a notable lack of systematic genomic research addressing the species' population, genetic background, differentiation patterns, and underlying drivers. This knowledge gap limits the development of genetically informed conservation actions and germplasm innovation, undermining the scientific rigor and long-term effectiveness of existing protection measures. Therefore, there is an urgent need for in-depth genetic studies to provide a robust basis for the effective conservation and sustainable utilization of *Primula beesiana* [7].

Compared with traditional RAD-seq, the principal technical innovation in ddRAD-seq lies in its capacity to generate a large number of single-nucleotide polymorphism (SNP) markers. This method includes a dual-restriction-enzyme strategy, which eliminates the uncertainties associated with random DNA fragmentation [8]. By utilizing two restriction endonucleases with distinct cleavage specificities, ddRAD-seq yields a more uniform and reproducible set of fragments, thereby significantly reducing sequencing bias. Moreover, as the technique does not require a reference genome, it is particularly well suited to studying non-model organisms [4]. In recent years, ddRAD-seq has been widely employed to assess genetic diversity and population structure in plants, uncovering patterns of genetic variation and differentiation across numerous species. These findings reflect the combined effects of ecological pressures and evolutionary history [9,10]. For instance, a study on *Cryptomeria japonica* var. sinensis Miq. revealed low genetic diversity and high differentiation among its seven populations, suggesting that historical glacial fluctuations and anthropogenic disturbances have shaped its genetic architecture [11]. Similarly, an analysis of 37 natural populations of caper (Capparis spinosa) identified six genetic lineages through ADMIXTURE analysis, illustrating a progressive decline in genetic diversity and effective population size [12]. Genomic studies on the endangered genus *Opisthopappus* demonstrated high genetic diversity in *Opisthopappus longilobus*, whereas *Opisthopappus taihangensis* showed signatures of recent genetic divergence, albeit with ongoing gene flow [13]. Correspondingly, research on 16 wild populations of *Strobilanthes biocullata* indicated low within-species genetic diversity and clear differentiation between southeastern and northwestern lineages, separated by the Xuefeng Mountains, a finding with important implications for breeding and conservation [14]. Collectively, these studies underscore the feasibility, effectiveness, and reliability of SNP markers developed via ddRAD-seq for population genetic analyses. This approach enables the detection of genetic diversity loss and extinction risk, thereby providing robust data support for biodiversity conservation and species management.

The application of ddRAD-seq in *Primula beesiana* allows for the clarification of inter-population genetic differentiation patterns, the identification of genetically vulnerable groups, and the quantification of gene flow dynamics. These insights provide critical data support for delineating evolutionarily significant units (ESUs) and formulating targeted strategies for in situ and ex situ conservation, as well as germplasm resource collection. In this way, this technique demonstrates substantial potential for enhancing the scientific foundation and precision of conservation strategies in this species. In this study, we collected 86 samples from six natural populations of *Primula beesiana* in Lijiang City, Yunnan Province, China. Using double-digest restriction site-associated DNA sequencing

(ddRAD-seq), we systematically evaluated genetic diversity, population structure, and gene flow through a series of analytical approaches—including genetic diversity indices, principal component analysis (PCA), population genetic structure inference, and gene flow estimation. This integrated approach revealed the key factors shaping genetic diversity and population structure in *Primula beesiana*. This study breaks away from the traditional research paradigm in this species, which has long relied on morphological observation. For the first time, it systematically analyses the genetic patterns of wild populations of *Primula beesiana*, an ornamental plant endemic to southwestern China, at the genomic level. This addresses a significant gap in our understanding of the genetic background of this horticulturally important species and facilitates the transition of horticultural practice from empirically based approaches to precision design and management informed by genetic data. Furthermore, through this research, by accurately identifying populations with high genetic diversity and formulating science-based introduction and ex situ conservation strategies according to population genetic differentiation patterns, we can effectively enhance cultivated adaptability and landscape performance, while reducing ecological and economic risks, in practical horticulture. This work yields direct scientific evidence supporting systematic conservation and sustainable horticultural development in this species.

2. Materials and Methods

2.1. Sample Collection

Field surveys conducted in the Lijiang City area of Yunnan Province, China, identified natural populations of *Primula beesiana* in six locations: Wenhai Village, Nanxi Village, Xuehua Village 1, Xuehua Village 2, Heishui Ercun, and Mudijing Reservoir. A total of 86 experimental specimens representing these six distinct natural populations were collected (Table 1, Figure 2). Healthy young leaves free from pests and diseases were selected as the experimental materials. Samples were assigned unique identifiers, placed in separate zip-lock bags with corresponding labels, and preserved in silica gel for subsequent DNA extraction. Sampling at each site followed these standardized protocols: (1) sample size was determined based on population size and plant density, with a minimum of five individuals collected per population; (2) a minimum spacing of 10 m was maintained between sampled individuals to avoid collecting closely related plants; (3) fresh, healthy mature leaves were collected from each individual, properly labeled, and immediately preserved in silica gel desiccant; (4) detailed records were made for each population, including the locality name, geographic coordinates (latitude and longitude), sample size, and elevation.

Table 1. Geographic information and sample sizes of six natural populations of *Primula beesiana* around Lijiang.

Location	County and Township	Population	Sample Size (n)	Longitude	Latitude	Elevation (m)
Wenhai Village	Baisha Town, Yulong County, Lijiang City, Yunnan Province, China	WH	16	100.1749° E	26.9754° N	3131
Nanxi Village	Huangshan Town, Yulong County, Lijiang City, Yunnan Province, China	NX	15	100.1567° E	26.7686° N	3093
Xuehua Village 1	Daju Township, Yulong County, Lijiang City, Yunnan Province, China	XHC1	7	100.2693° E	27.1906° N	3224

Table 1. *Cont.*

Location	County and Township	Population	Sample Size (n)	Longitude	Latitude	Elevation (m)
Xuehua Village 2	Daju Township, Yulong County, Lijiang City, Yunnan Province, China	XHC2	16	100.2501° E	27.1367° N	3025
Heishui Ercun	Daju Township, Yulong County, Lijiang City, Yunnan Province, China	HS	16	100.7250° E	27.5410° N	3207
Mudijing Reservoir	Daxing Town, Ninglang County, Lijiang City, Yunnan Province, China	MDJ	16	100.6750° E	27.6683° N	2876

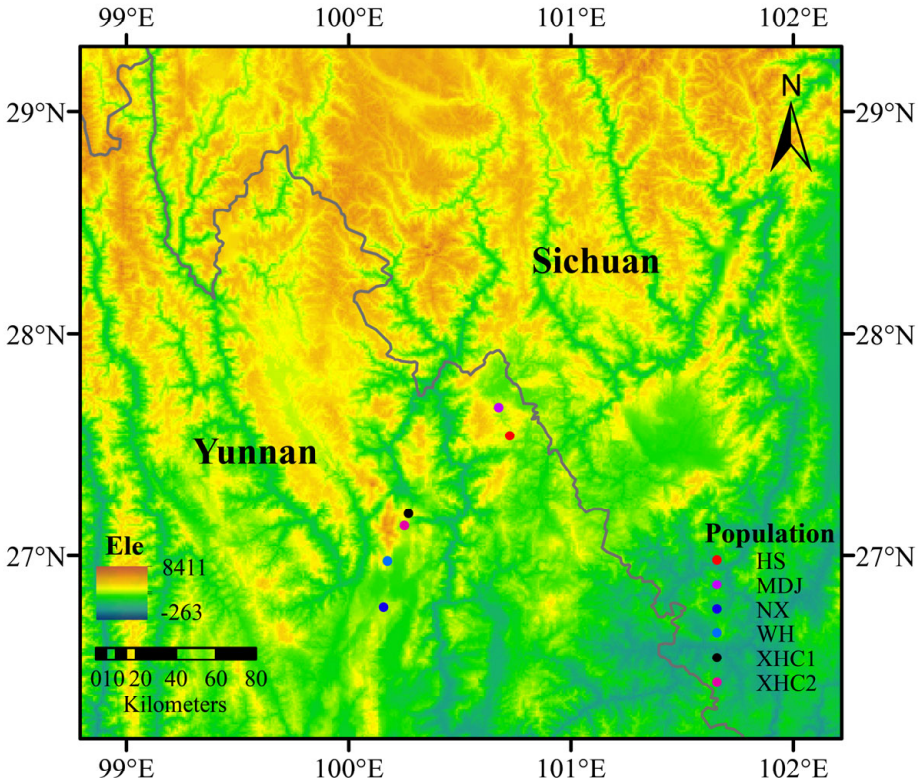


Figure 2. Distribution of six sampling sites for *Primula beesiana*.

2.2. DNA Extraction

Genomic DNA was isolated from healthy leaf tissue of *Primula beesiana* using a modified cetyltrimethylammonium bromide (CTAB) protocol [15].

2.3. ddRAD-seq Library Preparation and Sequencing

The ddRAD library was constructed according to the method described by Peterson et al., with modifications [8]. High-quality genomic DNA was digested with the restriction enzymes EcoRI and MseI in a reaction incubated at 37 °C for 8 h, followed by enzyme inactivation at 65 °C for 20 min. The digested DNA was stored at 12 °C for subsequent steps (Table 2). Digestion efficiency was verified using 1.5% agarose gel electrophoresis with 5 µL of the product. Enzyme digestion/ligation products from distinctly barcoded samples were pooled in equal volumes to ensure equimolar representation in the pooled library [8]. DNA fragments were separated via 2% agarose gel electrophoresis, and the target fragments (350–550 bp) were excised and purified using a gel extraction kit (Omega Bio-Tek, Norcross, GA, USA). The purified DNA fragments were used as templates for PCR amplification

in a 50 µL reaction system containing 10 × PCR buffer, dNTPs, adapter-specific primers, high-fidelity DNA polymerase, and template DNA. The PCR amplification procedure is as follows: first, pre-denature at 95 °C for 3 min. Then, enter the cycling phase, which consists of 15 cycles: denature at 95 °C for 30 s, anneal at 58 °C for 30 s, and extend at 72 °C for 30 s. After completing the cycles, extend at 72 °C for 5 min. Finally, store at 4 °C. The amplified products were analyzed using an Agilent 2100 Bioanalyzer (Santa Clara, CA, USA) to verify that fragment sizes were distributed within the 350–550 bp range and that concentrations met sequencing requirements. Then, library concentrations were accurately quantified using quantitative PCR (qPCR). Libraries passing quality control were subjected to paired-end 150 bp (PE150) sequencing on an Illumina NovaSeq platform, performed by a commercial service provider. To ensure sufficient sequencing depth, the expected data output was set at approximately 0.5 GB per sample, yielding a total data volume of about 42 GB.

Table 2. Double-digestion reaction system.

Components	Volume
High-quality genome DNA (200 ng, 20 ng/µL)	10 µL
Double-enzyme digestion mixture (10 U *EcoR* I and 10 U *Mse* I, NEB)	10 µL
10× enzyme digestion buffer	
Total volume	20 µL

2.4. Quality Control and Filtering of Raw Data

Following sequencing on the Illumina HiSeq PE150 platform, raw RAD-seq data were subjected to quality control using FastQC software (version 0.12.1) [16]. This tool generated comprehensive quality reports by evaluating key metrics such as data volume, quality score distribution, base composition, and GC content before and after filtering. The quality control process effectively removed low-quality reads, adapter contamination, and excessively short sequences, ensuring that the resulting clean reads met quality standards for subsequent analyses.

2.5. SNP Identification and Genotype Determination

The ddRAD-seq data were processed using the Stacks software package (version 2.70; http://catchenlab.life.illinois.edu/stacks/ (accessed on 27 March 2025) [17] following a structured analytical workflow: (1) Data Quality Control: Raw sequencing data were initially processed using the process radtags module in Stacks to perform demultiplexing and quality filtering. The parameter settings were as follows: process_radtags -1 xxx.R1.fq.gz -2 xxx.R2.fq.gz -b barcode.txt --renz-1 ecoRI --renz-2 mseI -c -q -r -o pro_out --len_limit 140 -t 135. (2) Locus Assembly per Sample: The ustacks module was used to assemble RAD loci for each sample based on the R1 read data. The parameter settings were M (number of mismatches allowed between alleles in a heterozygous locus) = 5 and n (minimum stack depth) = 2, following recommendations from Catchen et al. [17], and were validated in preliminary experiments. Each sample was assigned a unique identifier during this step. (3) Catalog Construction: A catalog of consensus loci was generated across all individuals using the cstacks module. (4) Locus Matching across Samples: The sstacks module was employed to match RAD tags from each sample against the catalog, producing a matches file that recorded polymorphism and sequencing depth information. The average per-locus sequencing depth ranged from 4× to 11×. (5) Data Format Conversion: The tsv2bam and gstacks modules were used to convert file formats and assemble paired-end reads in preparation for SNP calling. (6) SNP Calling and Filtering: SNP calling was performed

using the populations module [18]. The parameter M was set to 5 (the mismatch threshold for heterozygous genotypes), and n was set to 2 (the minimum stack depth). Sites with low frequency or excessive heterozygosity were filtered out. Summary statistics were exported to the populations.sumstats_summary.tsv file. (7) Output and Site Selection: All analysis results were saved in the bch_allsingleoutput directory. SNPs were categorized into selected and non-selected sets based on Tajima's D values [19]. For downstream population structure and principal component analyses, the first SNP from each locus was retained.

2.6. SNP Site Analysis

Statistical analysis of mutation types at identified SNP loci was conducted to characterize SNP distribution patterns and assess quality [20,21]. (1) Mutation type statistics: Python (version 3.13.1) scripts were employed to read VCF files, parsing the reference (Ref) and alternative (Alt) bases at each SNP site to quantify the 12 possible single-base substitution types. Classify substitution types as transitions (Ti) and transversions (Tv), separately quantifying their respective frequencies. Simultaneously, visualize the statistical results by plotting mutation type distribution charts and Ti/Tv ratio diagrams. (2) Calculate Transition/Transversion Ratio (Ti/Tv): Perform mutation type statistics and Ti/Tv ratio analysis on the obtained SNP sites. Employ Python scripts to count the 12 nucleotide substitution types, compute the frequencies of transitions (Ti) and transversions (Tv), calculate the Ti/Tv ratio, and evaluate sequencing data quality and mutation characteristics.

2.7. Genetic Diversity Analysis

To evaluate the genetic diversity of *Primula beesiana*, a multi-parameter analysis was performed on SNP data using GenAlEx 6.5 [22]. SNP genotypes were organized in Excel format according to individual and population identifiers to meet GenAlEx input specifications. The following genetic diversity indices were calculated: Number of Alleles (*Na*): The total number of alleles per locus, indicating the degree of gene polymorphism. Effective Number of Alleles (*Ne*): The number of alleles effectively contributing to genetic variation, calculated as $Ne = \frac{1}{\sum pi2}$, where pi is the frequency of the i-th allele. Shannon Diversity Index (*I*): A measure of allele richness and evenness, calculated as $I = -\sum (pi \times \ln pi)$. Observed Heterozygosity (*Ho*) and Expected Heterozygosity (*He*): Ho refers to the proportion of heterozygous individuals observed in the population, while *He* represents the expected heterozygosity under Hardy–Weinberg equilibrium [23], calculated as $He = 1 - \sum pi\hat{}2$. Nei's Gene Diversity (*H*): This reflects the overall level of genetic variation within a population, with values approximating *He*. Percentage of Polymorphic Sites (*PPB*): The proportion of polymorphic loci among all loci examined, serving as a comprehensive indicator of genetic diversity. Subsequently, one-way analysis of variance (ANOVA) was conducted using SPSS statistical software (version 31) to compare genetic diversity levels among populations, with statistical significance set at $p < 0.05$. This analysis allowed the systematic identification of populations exhibiting relatively high or low genetic diversity.

2.8. Analysis of Genetic Distance and Similarity Between Populations

To evaluate the degree of genetic differentiation and relatedness among populations, the following analyses were performed using SNP data: (1) Genetic Similarity and Distance Calculation: Nei's genetic distance (*D*) and genetic similarity coefficient (*I*) between populations were computed using GenAlEx 6.5, following Nei (1972) [24]. The genetic similarity coefficient was calculated as $I = \sum (p_{xi} \times p_{yi})$, where p_{xi} and p_{yi} represent the frequencies of the i-th allele in populations X and Y, respectively. Subsequently, the genetic distance was calculated using the formula $D = -\ln(I)$, where ln denotes the natural logarithm. (2) Visualization and Heatmap Generation: The resulting genetic distance and similarity matrices were imported into Python. Heatmaps were generated using the

Seaborn visualization library (https://seaborn.pydata.org/ (accessed on 27 January 2026)) to intuitively illustrate genetic relationships among populations.

2.9. Cluster Analysis

To investigate population clustering patterns and genetic structure in *Primula beesiana*, hierarchical clustering was performed based on Nei's genetic distance matrix. Analysis was conducted in Python using the hierarchy module from the SciPy library. Clustering was carried out under the complete linkage criterion, in which the distance between two clusters is defined as the maximum distance between any individual from one cluster and any individual from the other, a conservative approach suitable for revealing robust genetic relationships. A clustering threshold of 0.03 was applied to ensure biologically meaningful and stable groupings. The resulting dendrogram was visualized using the matplotlib library.

2.10. Principal Component Analysis

Principal component analysis (PCA) was performed on SNP data using the adegenet package in R (version 4.4.3) [25] to examine genetic distribution patterns among individuals and assess population genetic differentiation. SNP genotype data were imported into the R environment and converted into a genind object compatible with the adegenet package. Loci or individuals with high rates of missing data were filtered out to minimize potential bias in the analysis. PCA was conducted using the dudi.pca function to extract major axes of genetic variation. The first two principal components (PC1 and PC2), which captured the majority of genetic variance, were selected for visualization. Scatter plots of PC1 versus PC2 were generated using the ggplot2 visualization system, with individuals from different populations distinguished by color and symbol. Ninety-five percent confidence ellipses were superimposed to represent the distribution range of individuals within each population. The distribution and clustering patterns of points in the principal component space provided preliminary insights into the extent of genetic differentiation and potential genetic admixture among populations.

2.11. Mantel Test and Isolation-Based Distance (IBD) Analysis

To examine the relationship between genetic differentiation and geographical distance in *Primula beesiana*, a Mantel test was conducted within the isolation-by-distance (IBD) framework [26]. The analysis comprised the following steps: (1) Genetic Distance Matrix Construction: Pairwise Fst values between populations were computed using GenAlEx 6.5 to generate a genetic distance matrix. (2) Geographic Distance Matrix Calculation: Based on the latitude and longitude coordinates of each sampling site, straight-line geographical distances between populations were measured using ArcGIS 10.2, forming the geographic distance matrix. (3) Mantel Test Implementation: The Mantel test module in GenAlEx 6.5 was used to assess the correlation between the genetic and geographic distance matrices. The test was run with 9999 permutations to compute the Pearson correlation coefficient (r-value) and its statistical significance (p-value), evaluating whether genetic distance increased with geographic distance as predicted by the IBD model. (4) Result Visualization: A scatter plot of pairwise Fst values against geographic distances was generated in Python using the matplotlib and seaborn libraries, with a regression line fitted to illustrate the relationship.

2.12. Population Gene Flow Analysis

To quantify and visualize gene flow levels among geographical populations, genetic data were processed and analyzed using GenAlEx 6.5. Population-specific genetic differentiation coefficients (Fst) were calculated and arranged into symmetric matrices. The

gene flow (Nm) between each population pair was then estimated using Wright's formula under the neutral selection hypothesis: $Nm = \frac{1-FST}{4\times FST}$. This formula reflects the degree to which gene exchange counteracts genetic differentiation between populations. For visualization, the Nm values were represented as a matrix heatmap using Python with the matplotlib and numpy libraries. The color intensity in the heatmap corresponds to the relative strength of gene flow, providing a two-dimensional representation of inter-population gene flow patterns.

2.13. Analysis of Population Genetic Structure

To identify genetically distinct populations and assign individual membership probabilities, Bayesian clustering analysis was implemented in SNP data using STRUCTURE v2.3.4 [27]. (1) Parameter Configuration: The number of assumed genetic clusters (K) was varied from 1 to 7 to encompass the full range of possible population structures. An admixture model with correlated allele frequencies was applied to account for potential historical gene flow and shared allelic variation. Each run included a burn-in period of 100,000 steps followed by 100,000 Markov chain Monte Carlo (MCMC) iterations. For each value of K, 10 independent replicates were performed to assess clustering consistency. (2) Determination of Optimal K: The log-likelihood of the data, LnP(D), was recorded for each run. The optimal number of clusters was inferred using the ΔK method described by Evanno et al. [28] via the Structure Harvester web platform, where the value of K corresponding to the peak in the ΔK plot was selected. To integrate results across replicates for the selected K, CLUMPP [29] was used to align membership coefficients and address label switching and multimodality. Final membership probabilities were visualized using Distruct [30] to display individual assignments across inferred genetic clusters.

2.14. Analysis of Genetic Distance Between Individuals

To examine genetic relationships within and among geographical populations of *Primula beesiana* at the individual level, a genetic distance-based clustering analysis was conducted using MEGA X (version 10.2). The following steps were performed: (1) Data Preparation: SNP genotype data were converted into the FASTA format, compatible with MEGA X, with each individual represented as a nucleotide sequence. The resulting alignment was imported into MEGA X for preliminary validation, including checks for missing data and the exclusion of uninformative sites. (2) Genetic Distance Calculation: A pairwise genetic distance matrix was computed using the p-distance model. Gaps and missing data were handled through partial deletion to minimize their influence on genetic distance estimates. (3) Tree Construction: A dendrogram was generated from the genetic distance matrix using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm in the Phylogeny module. Branch support was assessed with 1000 bootstrap replicates. (4) Visualization and Interpretation: The resulting tree was visualized and annotated using the built-in tree visualization tools in MEGA X. Individual population affiliations were indicated, and topological patterns such as population-specific branching and admixture were examined. These results were later compared with those from PCA and STRUCTURE analyses to evaluate concordance across methods.

3. Results

3.1. SNP Site Statistics

Following site clustering and stringent quality filtering across all samples, a high-quality dataset comprising 1537 SNP loci was obtained. Analysis of nucleotide substitutions revealed twelve distinct mutation types (Figure 3). C > T and G > A transitions were the most frequent, occurring 372 and 314 times, respectively, and collectively accounting for

44.63% of all SNPs. In contrast, the T > G transversion was the least frequent, observed only 33 times. Among all SNPs, transitions (C > T, T > C, A > G, G > A) totaled 494, representing 32.14% of mutations and indicating considerable substitution bias between bases of the same type. Notably, C > T substitutions constituted 75.30% of all transitions. Transversions (e.g., C > A, G > T, A > T) were more prevalent, with 1043 occurrences accounting for 67.86% of SNPs. C > A and G > T transversions were relatively common, occurring 149 and 115 times, respectively. The overall distribution pattern suggests a noticeable transversion bias among SNP loci in the studied populations. As summarized in Figure 4, A total of 494 transition mutations and 1043 transversion mutations were detected, accounting for 32.1% and 67.9% of the total mutations, respectively. The transition-to-transversion ratio (Ti/Tv) calculated from these figures was 0.47.

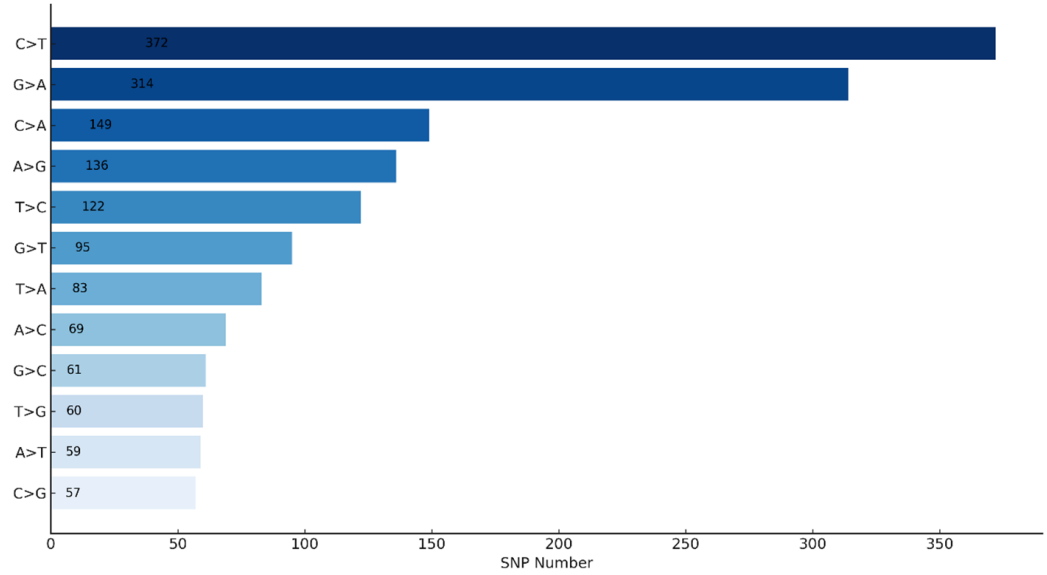


Figure 3. Mutation types and SNP counts.

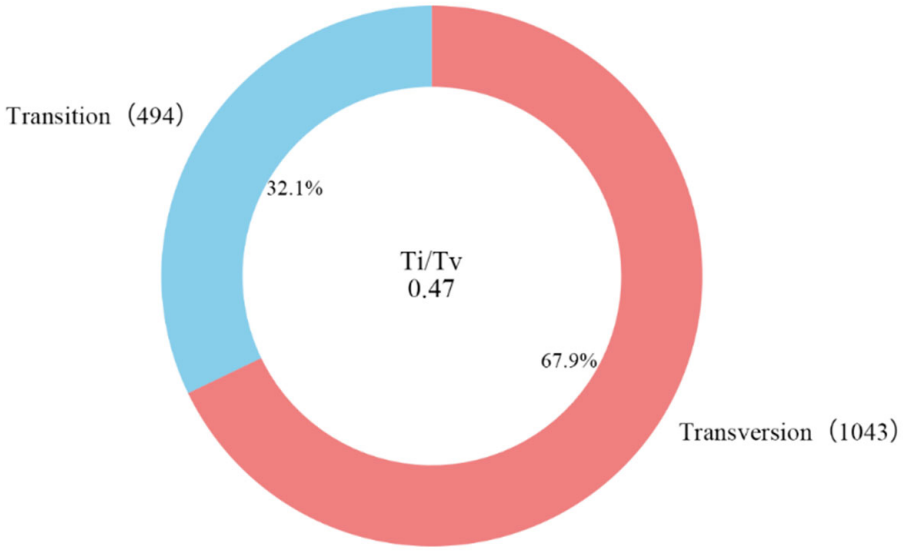


Figure 4. Transition vs. transversion mutation ratio (Ti/Tv).

3.2. Population Genetic Diversity

Analysis of genetic diversity across six geographical populations of *Primula beesiana* revealed distinct variation patterns (Table 3). The number of alleles (Na) across populations

decreased in the order HS (2.1516) > XHC2 (2.1509) > NX (2.1269) > WH (2.1093) > MDJ (2.0644) > XHC1 (1.7886), with a mean value of 2.0653. The effective number of alleles (Ne) ranged from 1.4108 (XHC1) to 1.4641 (HS), averaging 1.4365. Shannon's diversity index (I) varied between 0.4645 (XHC1) and 0.5001 (HS), with populations ordered as HS > NX > XHC2 > WH > MDJ > XHC1 and an overall mean of 0.4819. Regarding heterozygosity, the observed values (Ho) ranged from 0.0759 (XHC1) to 0.0849 (HS), while the expected heterozygosity (He) varied between 0.1068 (HS) and 0.1169 (WH). The mean Ho and He across all populations were 0.0803 and 0.1083, respectively. All populations showed consistently higher expected heterozygosity than that observed, indicating a general heterozygote deficit. Nei's gene diversity index (H) averaged 0.9882 across the six populations, with the highest value in XHC1 (0.9911) and the lowest in HS (0.9863). Statistical comparisons showed that populations NX, WH, and XHC2 consistently exhibited higher values across all six genetic parameters compared to HS, MDJ, and XHC1. The percentage of polymorphic loci (PPB) ranged from 31.24% (XHC1) to 49.97% (WH), with the following values: WH (49.97%), NX (47.82%), XHC2 (47.27%), HS (47.07%), MDJ (45.47%), and XHC1 (31.24%). In summary, all six populations of *Primula beesiana* maintained relatively high levels of genetic diversity, though significant inter-population differentiation was observed.

Table 3. Genetic diversity information for *Primula beesiana* provenances.

Population	Sample Size (N)	Number of Alleles (*Na*)	Effective Number of Alleles (*Ne*)	Shannon's Diversity Index (*I*)	Observed Heterozygosity (*Ho*)	Expected Heterozygosity (*He*)	Nei's Gene Diversity (*H*)	Number of Polymorphic Loci	Percentage of Polymorphic Loci (PPB%)
Wenhai Village (WH)	16	2.1093	1.4359	0.4816	0.0802	0.1169	0.9863	768	49.9675
Nanxi Village (NX)	15	2.1269	1.4459	0.4882	0.0819	0.1108	0.9877	735	47.8204
Xuehua Village 1 (XHC1)	7	1.7886	1.4108	0.4645	0.0759	0.0944	0.9911	477	31.2377
Xuehua Village 2 (XHC2)	16	2.1509	1.4286	0.4767	0.079	0.1078	0.9884	726	47.2656
Heishui Ercun (HS)	16	2.1516	1.4641	0.5001	0.0849	0.1068	0.9886	723	47.0703
Xuehua Village 2 (MDJ)	16	2.0644	1.4339	0.4802	0.0799	0.1130	0.9872	698	45.4723
Average Value	14.333	2.0653	1.4365	0.4819	0.0803	0.1083	0.9882	687.8333	44.8056

As illustrated in Figure 5, the genetic similarity coefficients among the six populations ranged from 0.8779 to 0.9481, with a mean of 0.9152, while pairwise genetic distances varied between 0.0533 and 0.1302, averaging 0.0889. These metrics collectively indicate generally close genetic relationships among all populations. Notably, populations XHC2 and WH showed the highest genetic similarity (0.9481) and smallest genetic distance (0.0533), suggesting a particularly close phylogenetic relationship and minimal genetic divergence. In contrast, populations XHC1 and MDJ exhibited the lowest genetic similarity (0.8779) and largest genetic distance (0.1302), indicating that they are the most distantly related among all populations, with comparatively greater genetic divergence.

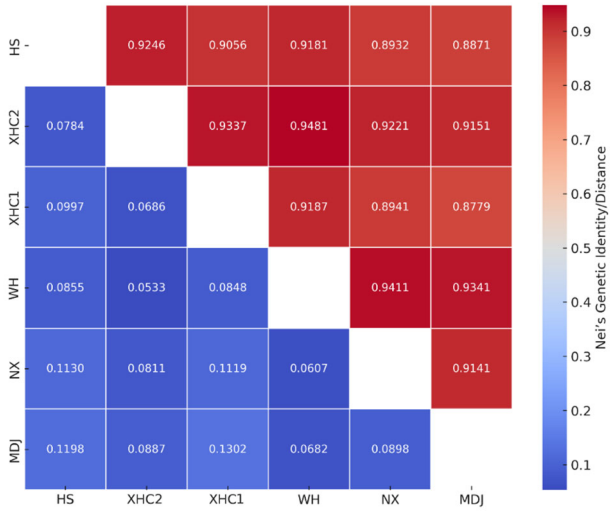


Figure 5. Heatmap analysis of Nei's genetic similarity (above the diagonal) and genetic distance (below the diagonal) among *Primula beesiana* populations. Note: The red section in the upper right represents Nei's genetic similarity, while the blue section in the lower left represents genetic distance.

3.3. Cluster Analysis of Primula beesiana

Figure 6 displays a dendrogram generated via complete-linkage hierarchical clustering of pairwise Fst values among the six populations of *Primula beesiana*, using a clustering threshold of 0.03. The six populations are grouped into four major clusters (from bottom to top in the dendrogram): the first cluster consists solely of MDJ, reflecting its distinct genetic identity; the second cluster includes NX, XHC2, and XHC1, indicating the high genetic similarity among them; and the third and fourth clusters contain HS and WH, respectively, with each forming an independent group with strong genetic distinctiveness. These clustering results are largely consistent with the genetic similarity coefficients shown in Figure 5. For example, XHC2 and XHC1, which cluster together, exhibit a high genetic similarity coefficient of 0.9337, confirming their close genetic relationship. In contrast, WH and MDJ, which show the lowest genetic similarity (0.8779) and largest genetic distance (0.1302), occupy the most distant positions in the dendrogram, underscoring their pronounced genetic differentiation.

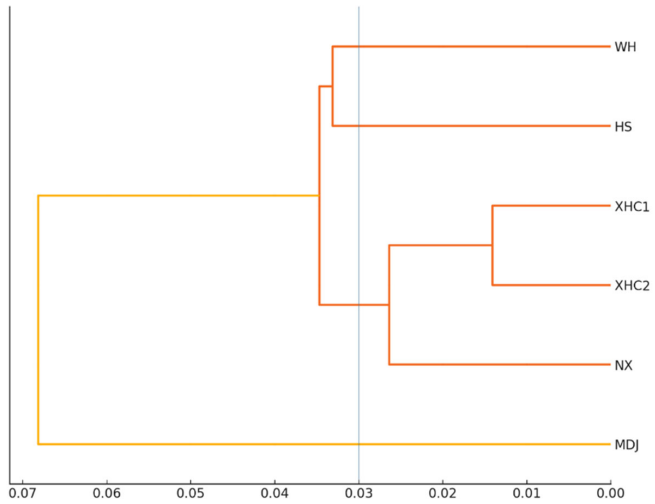


Figure 6. Cluster analysis of 6 provenances in *Primula beesiana*.

3.4. Principal Component Analysis of Primula beesiana

Figure 7 presents the principal component analysis (PCA) results for 86 *Primula beesiana* individuals based on SNP data. In the PCA scatter plot, individuals of the MDJ population clustered in the right-hand quadrants (first and fourth), with their confidence ellipse showing no overlap with those of other populations, indicating strong genetic distinctiveness. All individuals of the WH population were located in the third quadrant, also demonstrating clear genetic differentiation. Individuals of the HS population were widely distributed across the second and third quadrants, showing substantial genetic variation and clear separation from other populations. These patterns support the classification of MDJ, WH, and HS as three distinct genetic clusters. In contrast, individuals of the NX, XHC2, and XHC1 populations formed tightly overlapping clusters, with largely congruent confidence ellipses, reflecting high genetic similarity among these three groups.

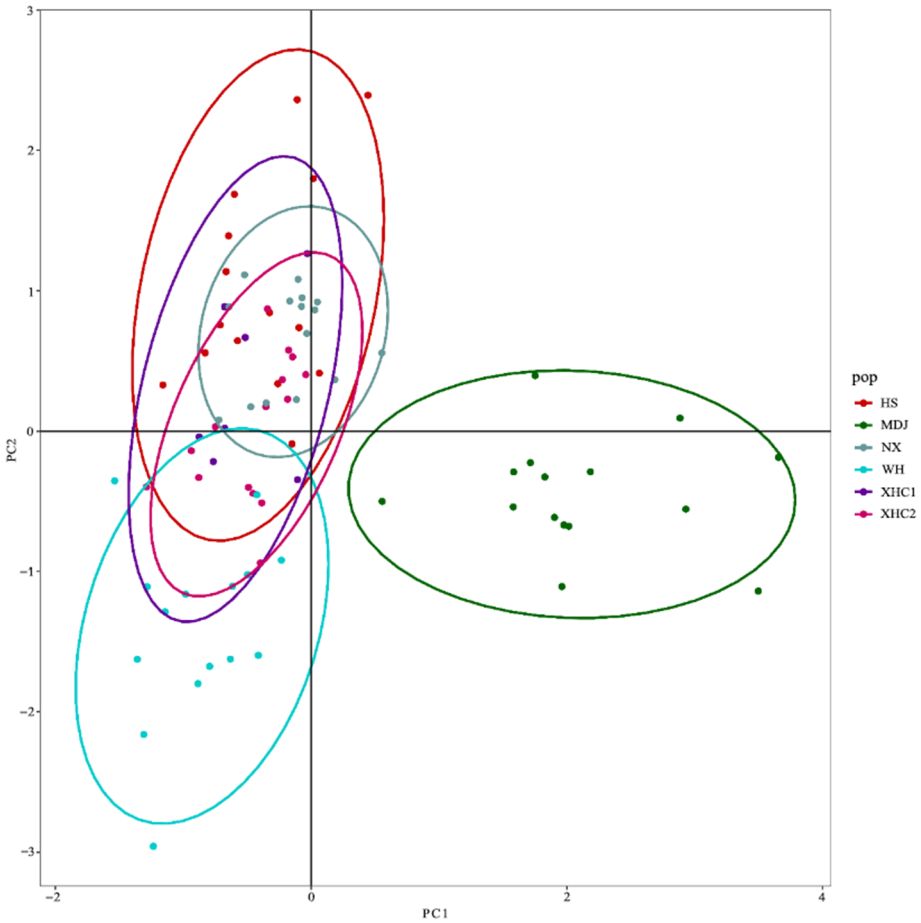


Figure 7. PCA scatter plot of *Primula beesiana* populations.

3.5. Mantel Test and Isolation-Based Distance (IBD) Model Analysis

A Mantel test revealed a strong positive correlation between genetic differentiation (Fst) and geographical distance (r = 0.854, p = 0.001), indicating a significant pattern of isolation by distance (IBD) in *Primula beesiana* at the regional scale. Analysis of genetic diversity indices (Table 3) showed that the geographically isolated MDJ population exhibited a below-average effective allele number and Shannon's diversity index, whereas populations such as NX demonstrated higher diversity. This pattern suggests restricted gene flow in geographically isolated populations such as MDJ, leading to increased genetic differentiation. Pairwise comparisons of Fst values and geographical distances (Table 4) further supported this relationship. Despite being separated by 66.42 km, MDJ and XHC1

exhibited the highest Fst value (0.04068). Conversely, the geographically closest populations, XHC2 and XHC1 (6.27 km apart), showed the lowest Fst (0.00557), indicating more frequent gene flow over shorter distances. Notably, although NX and MDJ represent the most geographically distant pair (112.14 km apart), their Fst value (0.03338) was lower than that between MDJ and XHC1, suggesting that factors beyond geographical distance may modulate genetic differentiation at local scales. The scatter plot in Figure 8 visually confirms this positive correlation, showing Fst values increasing with geographical distance (range: 6.27–112.14 km). Cluster analysis (Figure 6) corroborates these findings, with MDJ forming a distinct cluster while the geographically proximate NX, XHC2, and XHC1 groups cluster together. Collectively, these results demonstrate that genetic differentiation in *Primula beesiana* generally follows an isolation-by-distance pattern, though local factors may modify the influence of geographical distance on gene flow.

Table 4. Integrated comparison of Fst (above the diagonal) and geographic distances (below the diagonal).

	HS	XHC2	XHC1	WH	NX	MDJ
HS	0	0.02025	0.02166	0.02317	0.0197	0.03417
XHC2	64.93	0	0.00557	0.01737	0.01329	0.03122
XHC1	59.51	6.27	0	0.02204	0.01715	0.04068
WH	83.04	19.36	25.61	0	0.01965	0.03803
NX	102.47	41.83	48.08	22.99	0	0.03338
MDJ	14.94	72.36	66.42	91.35	112.14	0

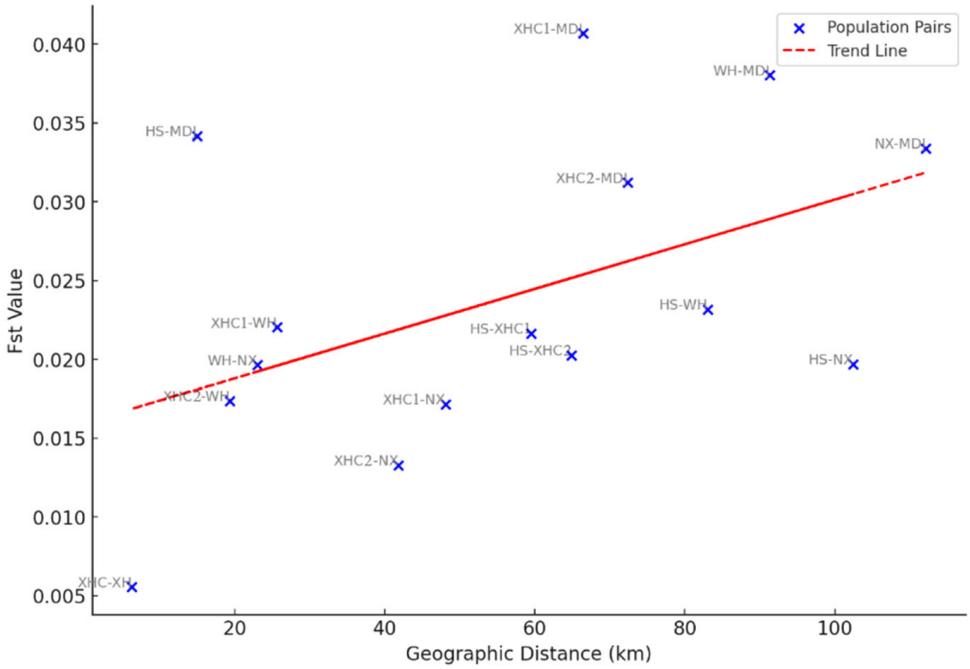


Figure 8. Mantel test results between Fst values and geographic distances among *Primula beesiana* populations.

3.6. Analysis of Gene Flow Between Populations

Figure 9 presents a heatmap of gene flow (Nm) between populations. The Nm values ranged from 5.90 to 44.69, indicating substantial spatial heterogeneity in gene flow intensity among populations of *Primula beesiana*. The strongest gene flow occurred between XHC2 and XHC1 (Nm = 44.69), which are separated by only 6.27 km and located near the same village. This suggests that spatial proximity and habitat similarity effectively promote genetic exchange. The second-highest gene flow was observed between

XHC2 and NX (Nm = 20.03); although separated by a greater distance, these populations maintain substantial genetic connectivity, potentially facilitated by ecological corridors or human-mediated dispersal. In contrast, gene flow involving the MDJ population was generally low. The lowest Nm value (5.90) was observed between MDJ and XHC1, corresponding to their considerable geographical separation (66.42 km) and distinct ecological conditions. MDJ is situated in an eastern reservoir wetland area characterized by lower elevation and enclosed topography, creating pronounced microenvironmental isolation from the high-altitude mountain slopes inhabited by XHC1. This demonstrates how natural terrain and environmental filtering jointly constrain genetic exchange, further supporting the genetic distinctiveness of MDJ. Overall, this study reveals significant variation in gene flow intensity among the six populations of *Primula beesiana*, with a clear correlation with geographical distance and ecological compatibility.

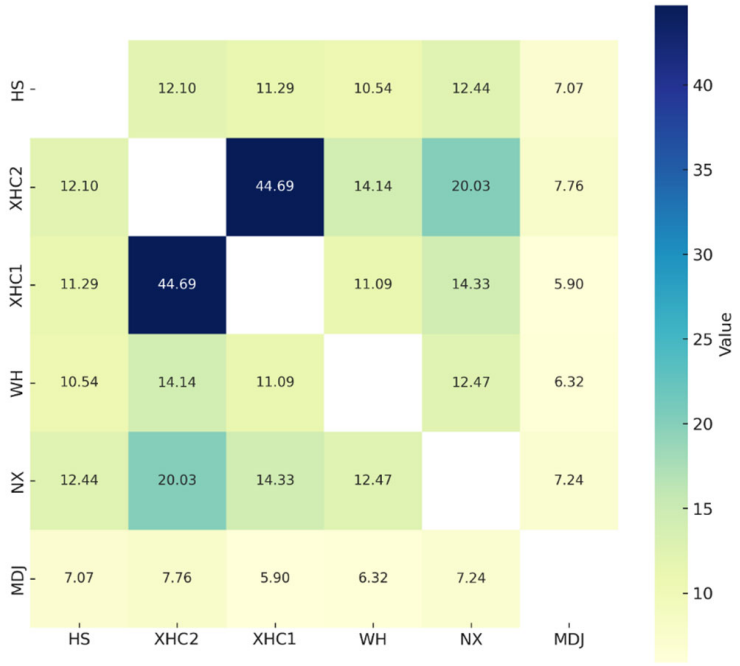


Figure 9. Heatmap of gene flow intensity matrix among populations.

3.7. Analysis of Population Genetic Structure in Primula beesiana

As shown in Figure 10a, the log probability of the data, LnP(D), increased significantly from K = 1 to K = 4, indicating a substantial improvement in model fit over this range. Beyond K = 4, the rate of increase in LnP(D) slowed considerably, and the inferred cluster structure exhibited high explanatory power and robustness. Therefore, K = 4 was identified as the optimal number of genetic clusters. Figure 10b shows the change in ΔK across different values of K. A distinct peak in ΔK was observed at K = 4, further supporting this as the optimal cluster number. This result is consistent with the earlier clustering analysis, indicating that the six populations of *Primula beesiana* can be grouped into four genetic clusters: the first comprising MDJ, the second HS, the third WH, and the fourth XHC1, XHC2, and NX. This grouping is also supported by the Q-matrix derived from STRUCTURE analysis. Figure 11 illustrates the genetic composition of the populations under K = 4. The MDJ population is characterized by a highly homogeneous green genetic component, consistent with its geographically isolated status. The WH population is dominated by a distinct red component, showing high internal uniformity and clear differentiation from other groups. The HS population is primarily represented by a cyan component, with a consistent secondary presence of purple, supporting its recognized status as an independent genetic cluster. In contrast, the XHC1, XHC2, and NX populations share a

mixed genetic composition dominated by cyan but with varying proportions of purple and red components, reflecting a pattern of genetic admixture suggestive of ongoing gene flow among these three populations.

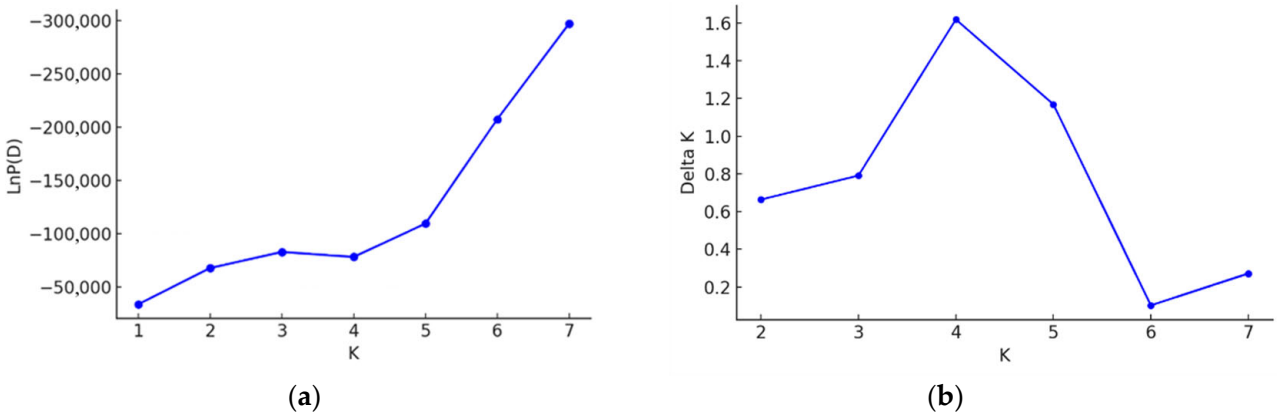


Figure 10. (**a**) Trend of LnP(D) with K value. (**b**) Trend of ΔK with K value.

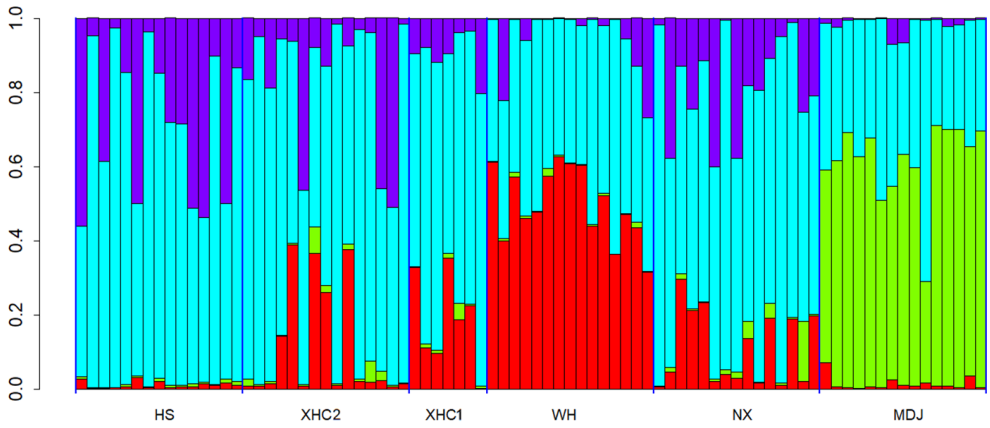


Figure 11. Group genetic structure of sample material.

3.8. Analysis of Genetic Distance Between Individuals in Primula beesiana

The dendrogram in Figure 12 (distance scale: 0.000–0.025) illustrates pronounced genetic divergence and a clear population structure among the studied populations. The MDJ population exhibits strong genetic distinctiveness, with most individuals (e.g., MDJ-1-MDJ-9, MDJ-12, MDJ-16) forming a well-defined cluster in the lower-middle section of the tree, clearly separated from other populations. This pattern suggests that the MDJ population may have experienced long-term geographical or ecological isolation, resulting in restricted gene flow. Similarly, the WH population demonstrates a clustering tendency, with the majority of its individuals (e.g., WH-3-WH-8, WH-12, WH-15) located in the upper-middle part of the tree and forming a relatively independent branch. The HS population (e.g., HS-2, HS-4, HS-8) is primarily distributed in the upper region of the dendrogram, also showing a concentrated distribution, further supporting the genetic independence of both WH and HS. In contrast, individuals from the XHC2, XHC1, and NX populations are extensively intermingled. For instance, several samples from XHC2 and XHC1 (e.g., XHC2-11 and XHC1-4, XHC2-5 and XHC1-1) are closely positioned in the upper part of the tree, indicating a close genetic relationship. A similar pattern is observed in the NX population, with individuals such as NX-13 and NX-7 showing an admixed distribution with XHC2 and XHC1, suggesting frequent gene flow among these three populations, likely

facilitated by geographical proximity or ecological similarity. High bootstrap values at key nodes further support the reliability of the inferred topological structure.

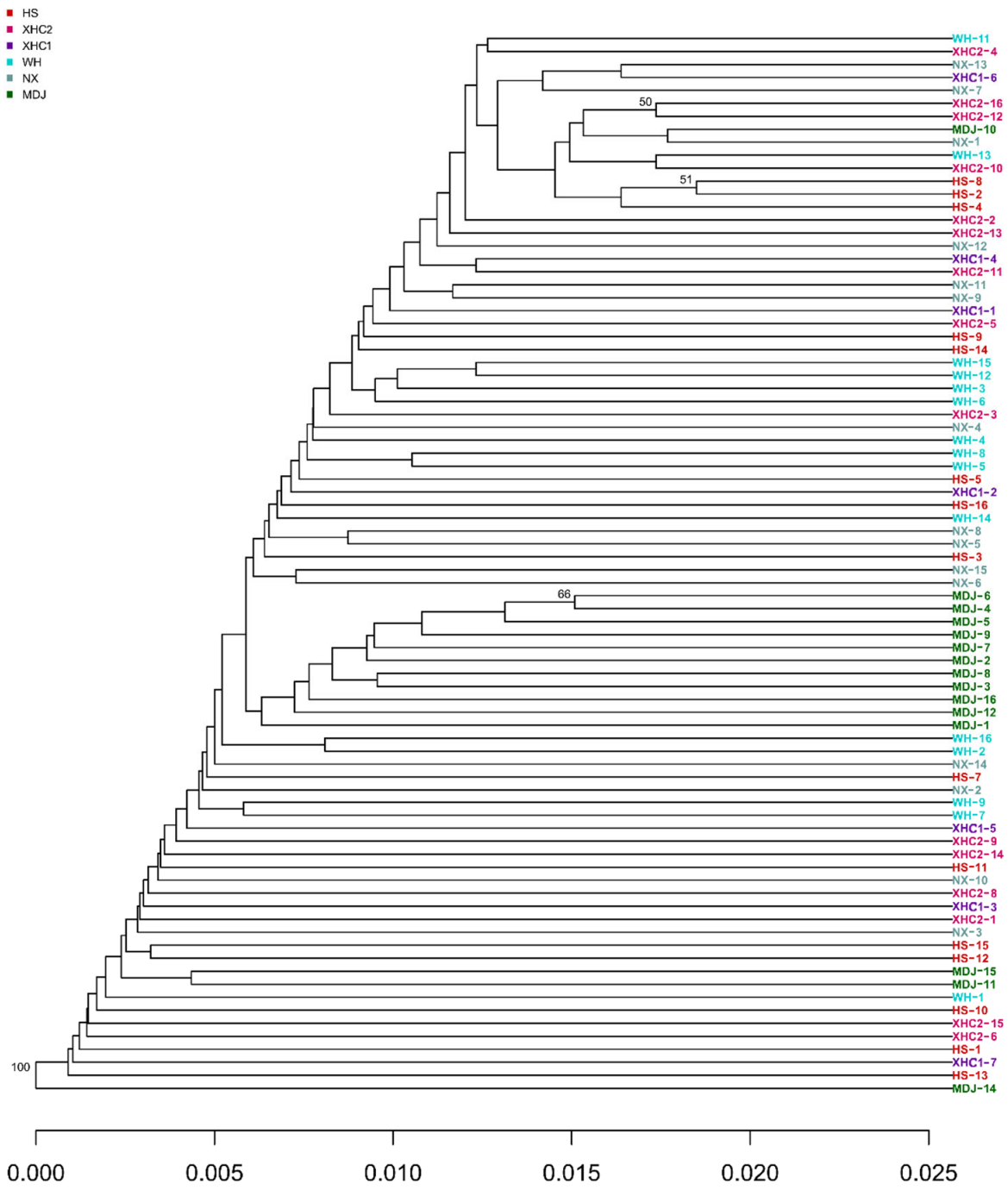


Figure 12. Dendrogram of genetic distance between individuals.

4. Discussion

The transition-to-transversion ratio (Ti/Tv) serves as a crucial indicator in evaluating genomic genetic variation and a key parameter in quality control in high-throughput sequencing data. A Ti/Tv ratio below the typical level observed in common species may be associated with environmental stresses unique to alpine wetlands, such as intense ultraviolet radiation [31], temperature fluctuations [32], and oxidative damage [33,34]. The distribution of mutation types also reflects the genome's response characteristics to environmental stresses. Under natural conditions, transition mutations are relatively stable,

whereas transversion mutations typically occur at lower frequencies. Consequently, the increased proportion of transversion mutations may indicate that *Primula beesiana* has undergone intense selective pressure within alpine wetland ecosystems. Further analysis suggests this mutational bias likely arises from the combined effects of persistent ultraviolet radiation and low-temperature stress in high-humidity, high-altitude environments. These conditions exacerbate DNA damage and repair processes, thereby promoting frequent transversions. This phenomenon has been corroborated in numerous studies. For instance, under elevated temperatures, oxidative damage leads to increased DNA mutation rates, particularly during specific cell cycle phases. Reports indicate that oxidative damage primarily impacts the G1 and S phases of the cell cycle, likely due to the reduced DNA repair capacity in cells during these stages. This represents a common adaptive response in plants at high altitudes [35,36]. In such environments, plants generally exhibit similar adaptive responses by increasing mutation frequency to enhance genomic plasticity. These mutations not only directly contribute to genetic diversity within populations but also provide a potential genomic foundation for the subsequent adaptive evolution of the species.

This pattern aligns with findings in European primrose (*Primula vulgaris*), where a clear genetic diversity gradient was observed, with populations in the eastern Caucasus and Carpathian Mountains displaying significantly higher allelic richness than those in western lowlands [37]. Similarly, in the Lijiang region of Yunnan, China, *Primula beesiana* exhibits notable spatial genetic differentiation: populations in contiguous habitats with sufficient pollinators and minimal anthropogenic disturbance (e.g., XHC2, WH) maintain high genetic diversity, whereas those in remote or topographically isolated areas (e.g., MDJ) experience limited gene flow. The MDJ population, situated at a relatively low elevation (2876 m) and naturally isolated by surrounding higher mountains, exemplifies how physical barriers can reduce genetic diversity. Thus, the coexistence of high intraspecific variation (over 97% of genetic variation is found within populations) alongside reduced diversity in specific regional populations reflects the complex evolutionary dynamics of *Primula beesiana* at fine geographical scales.

Genetic diversity arises from dynamic changes in allele number and frequency within a species, a process governed by multiple evolutionary mechanisms—including genetic drift, gene mutation, and reproductive modes—and further modulated by anthropogenic environmental disturbances [38]. As a key regulatory factor, plant mating systems shape the genetic architecture of populations by influencing heterozygosity and genetic differentiation, thereby affecting the spatiotemporal distribution of gene flow [39]. The evolutionary trade-off between outcrossing and selfing strategies is particularly significant. Theoretically, *Primula beesiana* could maintain high outcrossing rates via insect pollination. However, the average observed heterozygosity across the six studied populations was significantly lower than expected, with a heterozygote deficiency rate of approximately 25.9%. This overall heterozygosity deficit in *Primula beesiana* may result from habitat fragmentation or abrupt population decline, which reduces the effective mating pool and increases the likelihood of inbreeding. In particular, within the areas covered, certain populations are distributed across farmland and zones of frequent human activity, where environmental fragmentation is unavoidable. This may well be one reason for the overall low heterozygosity observed in *Primula beesiana*. From a conservation standpoint, the genetic vulnerability observed in the XHC1 and MDJ populations—attributable to small population sizes and restricted gene flow—requires urgent attention. This pattern is consistent with findings from numerous studies, particularly in small or isolated populations where a loss of genetic diversity is often linked to genetic drift and inbreeding depression [40–42].

Genetic differentiation, a fundamental concept in population genetics, quantifies the divergence in genetic composition among populations. Analysis of molecular variance

(AMOVA) revealed that approximately 97.62% of molecular variation in *Primula beesiana* occurs within populations, while only 2.38% is distributed among populations. This pattern aligns with genetic structures observed in other ecologically adaptable plant species, where high within-population diversity coexists with limited inter-population differentiation [43,44]. Previous studies on European *Primula vulgaris* populations similarly reported generally low Fst values, indicating maintained gene flow across geographical scales [37]. Consistently, the six geographical populations of *Primula beesiana* in this study exhibited overall low Fst values despite significant altitudinal and topographical variation, demonstrating how gene flow effectively counteracts genetic differentiation. Notably, the lowest Fst was observed between the geographically proximate XHC2 and XHC1 populations, where spatial proximity and habitat similarity likely facilitate pollinator movement and seed dispersal, thereby minimizing genetic divergence. In contrast, the highest Fst occurred between MDJ and XHC1, corresponding to their substantial geographical separation, with distance and topographical barriers collectively restricting gene flow. This observation corroborates studies on *Rhizophora* species, where molecular phylogenetics and population genetic analyses using multiple nuclear genes confirmed that geographical isolation imposes localized constraints on gene exchange [45].

Inter-population gene flow (Nm) serves as a crucial indicator in assessing the level of genetic exchange between populations. Higher gene flow typically implies lower genetic differentiation between populations, as gene exchange reduces genetic variation through mechanisms such as migration and diffusion [46]. When $Nm < 1$, genetic drift often exerts a greater influence on differentiation than gene flow; conversely, when $Nm > 1$, gene flow is sufficient to suppress inter-population genetic differentiation to a certain extent [5,47]. Analysis reveals that Nm values among the six populations are generally greater than 1. Moderate gene flow was detected among populations of *Primula beesiana*, with the strongest exchange occurring between XHC2 and XHC1. This pattern reflects their geographical proximity and comparable ecological conditions, which jointly facilitate pollen and seed dispersal. In contrast, gene flow was lowest between MDJ and XHC1, constrained by both topographic barriers and pronounced habitat heterogeneity. Similarly, genetic exchange between MDJ and both HS and XHC2 remained limited, indicating that the MDJ population is largely genetically isolated within the species' metapopulation system. The observed low genetic diversity and heterozygosity deficit in the MDJ population may result from long-term independent evolution and restricted gene flow, potentially due to historical bottleneck or founder effects [48–50]. These processes have likely intensified its genetic divergence from other populations.

Genetic differentiation (Fst) among the six populations of *Primula beesiana* showed a significant positive correlation with geographical distance, supporting the isolation-by-distance (IBD) model as a major driver of population structure in this species. However, certain population pairs—such as HS-MDJ and NX-XHC2—displayed levels of genetic differentiation not fully consistent with geographical proximity. For instance, despite the relatively short distance between HS and MDJ, their elevated Fst values suggest that factors such as topographic barriers, habitat heterogeneity, or historical divergence may locally override or amplify IBD effects. This phenomenon is supported by several previous studies. Research in a narrow hybrid zone in the Iberian Peninsula demonstrated that gene flow between two cryptic populations was substantially shaped by microenvironmental variation, leading to asymmetric genetic exchange and reinforcing isolation [51]. Similarly, populations of *Eperua falcata* Aubl. under contrasting environmental conditions exhibited pronounced genetic divergence, likely driven by natural selection [52]. Genetic structure analysis of the blue fan palm (*Brahea armata*) further revealed that although geographical isolation strongly influenced differentiation, long-distance gene flow persisted, particularly

across climatically favorable areas [53]. These patterns illustrate how plant populations can maintain genetic connectivity over long distances despite geographical separation—a phenomenon also observed in alpine plants of the Tibetan Plateau [54]. Moreover, studies on *Iris atropurpurea* revealed that ecological differences exerted a more pronounced influence on reproductive isolation than geographical distance, further underscoring the dominant role of environmental factors in population genetic differentiation processes [55]. Certain sampling sites within the study area were proximate to farmland, reservoirs, or tourist development zones. Human activities such as the deliberate relocation of seedlings, landscaping projects, and road construction may inadvertently facilitate pollen or seed dispersal between populations. This locally circumvents the constraints of geographical distance on gene flow, resulting in a more complex spatial genetic structure. The generally low Fst values and high Nm values observed in *Primula beesiana* indicate that gene exchange between populations remains unimpeded. This facilitates the maintenance of high intrapopulation genetic diversity in geographically adjacent areas and enhances resilience to environmental disturbances. However, certain local populations (e.g., MDJ, HS, XHC1) exhibit significant differentiation from neighboring populations due to historical, environmental, or anthropogenic factors. While this differentiation may foster adaptive gene pools within specific microhabitats, it also implies heightened genetic risk in these populations [56]. Research indicates that even at close geographical distances, strong selective pressures across different ecological gradients can drive genomic differentiation. This microgeographic selection pattern indicates that spatial environmental heterogeneity aids in maintaining intrapopulation genetic diversity [57]. Consequently, conservation management should treat these markedly differentiated local populations as Evolutionarily Significant Units (ESUs) warranting priority protection. Should these populations be lost due to external disturbances, the unique gene pools that they harbor would be irrecoverable, resulting in irreversible genetic loss. Consequently, conservation efforts for *Primula beesiana* must balance two primary objectives: maintaining overall gene flow and safeguarding distinctively differentiated populations. For instance, research indicates that gene flow is crucial in preserving species genetic diversity and adaptive capacity. However, simultaneously, distinctively differentiated populations may show unique adaptations to specific environments, necessitating special protection [58,59].

The six geographical populations of *Primula beesiana* can be classified into four major genetic clusters: MDJ, WH, and HS each form distinct groups, while NX, XHC1, and XHC2 cluster together. The results from principal component analysis (PCA) are largely consistent with the dendrogram topology. In the PCA scatter plot, MDJ individuals are concentrated on the right side, showing minimal overlap with other populations, reflecting a highly differentiated genetic background. Similarly, WH and HS form relatively independent clusters, whereas NX, XHC1, and XHC2 exhibit substantial overlap, indicating closer phylogenetic relationships or more frequent gene flow among them. These patterns align with findings from Fst and gene flow (Nm) analyses: geographically adjacent or ecologically similar populations (e.g., XHC2-XHC1-NX) maintain closer genetic connections, whereas relatively isolated populations such as MDJ demonstrate stronger genetic independence. This is consistent with the general principle that geographically or ecologically proximate populations tend to show reduced genetic divergence, as observed in Brazilian tall coconut populations, where genetic similarity decreases with geographical distance [60]. Within the overall genetic structure, the distinct differentiation of MDJ, WH, and HS merits further investigation to clarify the underlying ecological or historical causes. In the overall clustering results, the distinct differentiation of MDJ, WH, and HS populations to varying degrees corroborates the hypothesis that ‘geographical proximity coupled with ecological similarity’ may enhance gene flow between populations. Simultaneously, this highlights

the MDJ population's status as a 'genetic island' within the region. Combining analyses of the MDJ population's geographical location, microenvironmental differences, and potential historical bottlenecks, it can be inferred that its prolonged independent evolution has resulted in low genetic diversity and significant differentiation from other populations. Priority conservation strategies should therefore be implemented for this population.

Studies indicate that *Primula beesiana* exhibits considerable plasticity in environmental adaptation. These findings not only offer key insights into the evolutionary mechanisms and adaptability of this species but also highlight the importance of developing differentiated conservation strategies from the perspectives of germplasm preservation and population management. By rationally planning potential gene exchange pathways among distinct populations and implementing targeted conservation measures in key isolated populations, it is possible to maintain the species' high genetic diversity and long-term evolutionary potential across its distribution range. Specifically, maintaining ecological corridors and regulating land use can help preserve pollination pathways and seed dispersal routes between adjacent populations. Meanwhile, highly isolated populations with low genetic diversity (e.g., MDJ) should be prioritized for conservation and continuous monitoring to prevent inbreeding depression and excessive genetic drift, thereby sustaining regional-scale diversity and evolutionary potential [61,62]. Furthermore, based on the four major genetic clusters (MDJ, WH, HS, and a tightly clustered group comprising XHC2, XHC1, and NX) of *Primula beesiana* identified in this study, along with its gene flow patterns, we propose a systematic strategy, spanning from resource conservation to variety development. The core strategy involves utilizing the XHC2-XHC1-NX population—characterized by high gene flow and rich genetic diversity—as a foundational breeding germplasm repository to broadly aggregate adaptive genes. Specific breeding programs include the following: Breeding Broadly Adaptable Cultivars: Conducting preferential hybridization within the XHC2, XHC1, and NX populations to enhance their inherent adaptability. Synergistic Conservation and Breeding: Designating the three independent clusters MDJ, WH, and HS as Evolutionarily Significant Units (ESUs) for priority in situ conservation while simultaneously utilizing their germplasm as a strategic breeding resource and maintaining ecological corridors between adjacent populations such as XHC2-XHC1-NX to ensure sustained gene flow and natural evolutionary processes. This approach translates genetic information into precise parental selection and sustainable germplasm management, enabling the sustainable performance of both species conservation and horticultural practices.

5. Conclusions

Primula beesiana maintains relatively high overall genetic diversity yet exhibits significant inter-population differentiation. Populations such as WH, NX, and XHC2 display comparatively rich diversity, whereas MDJ and XHC1 show reduced levels, potentially due to factors including limited sample size, geographical isolation, and historical genetic drift. Genetic differentiation among populations is generally low, with most molecular variance occurring within populations. A significant positive correlation was observed between genetic differentiation and geographical distance ($r = 0.854$, $p = 0.001$), supporting the isolation-by-distance (IBD) model. However, local populations such as MDJ exhibit markedly elevated differentiation beyond that predicted by distance alone, likely reflecting the influence of topographical or ecological barriers. Gene flow analysis reveals frequent genetic exchange between geographically adjacent populations (e.g., XHC2 and XHC1), whereas spatially isolated populations (e.g., MDJ) show restricted connectivity. The six populations form four genetic clusters, with MDJ, WH, and HS constituting distinct groups recommended for priority conservation as independent evolutionary significant units (ESUs). In contrast, NX, XHC1, and XHC2 demonstrate strong local adaptation and dis-

persal potential. At the regional scale, *Primula beesiana* exhibits a genetic architecture characterized by high within-population diversity, limited inter-population differentiation, and pronounced local isolation in specific populations.

Author Contributions: Conceptualization, Q.L.; data curation, Y.M. and P.X.; methodology, Q.L. and Z.L.; formal analysis, S.Z. and Z.D.; writing—original draft, Q.L. and Z.L.; writing—review and editing, P.X.; project administration, Y.M. and P.X.; funding acquisition, Q.L. and P.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Key Technologies Research for the Germplasm of Important Woody Flowers in Yunnan Province (No. 202302AE090018); Yunnan Provincial Science and Technology Mission (No. 202404BI090014).

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 conflicts of interest.

References

1. Editorial Board of Flora of China. *Flora of China*; Science Press: Beijing, China, 1990; Volume 59, ISBN 9787030012869.
2. Liang, S. Introduction and Cross Breeding of Some Wild Primula in Southwest China. Ph.D. Thesis, Beijing Forestry University, Beijing, China, 2006.
3. Egan, P.A.; Price, M.F. *Mountain Ecosystem Services and Climate Change: A Global Overview of Potential Threats and Strategies for Adaptation*; UNESCO: Paris, France, 2017; ISBN 978-92-3-100225-0.
4. Andrews, K.R.; Good, J.M.; Miller, M.R.; Luikart, G.; Hohenlohe, P.A. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **2016**, *17*, 81–92. [CrossRef]
5. Frankham, R.; Ballou, J.D.; Briscoe, D.A. *Introduction to Conservation Genetics*; Cambridge University Press: Cambridge, UK, 2002; ISBN 9780521702713.
6. Lenore, F. Effects of habitat fragmentation on biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **2003**, *34*, 487–515. [CrossRef]
7. Primack, R.B. *Essentials of Conservation Biology*; Sinauer Associates Sunderland: Sunderland, UK, 2006; ISBN 9781605352893.
8. Peterson, B.K.; Weber, J.N.; Kay, E.H.; Fisher, H.S.; Hoekstra, H.E. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **2012**, *7*, e37135. [CrossRef] [PubMed]
9. Etter, P.D.; Bassham, S.; Hohenlohe, P.A.; Johnson, E.A.; Cresko, W.A. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Mol. Methods Evol. Genet.* **2011**, *772*, 157–178. [CrossRef]
10. Zhang, S.; Kang, H.; Yang, W. Population genetic analysis of *Nyssa yunnanensis* by reduced-representation sequencing technique. *Bull. Bot. Res.* **2019**, *39*, 899–907. [CrossRef]
11. Cai, M.; Wen, Y.; Uchiyama, K.; Uchiyama, K.; Onuma, Y.; Tsumura, Y. Population genetic diversity and structure of ancient tree populations of *Cryptomeria japonica* var. sinensis based on RAD-seq data. *Forests* **2020**, *11*, 1192. [CrossRef]
12. Wang, Q.; Zhang, H. Population genetic structure and biodiversity conservation of a relict and medicinal subshrub Capparis spinosa in arid Central Asia. *Diversity* **2022**, *14*, 146–162. [CrossRef]
13. Hou, H. Population Genetic Analysis of *Opisthopappus Shih*-Based on Simplified Genomic Data. Master's Thesis, Shanxi Normal University, Taiyuan, China, 2020. [CrossRef]
14. Wang, Q. Study on Genetic Diversity of *Baphicacanthus cusia*. Master's Thesis, Nanjing Forestry University, Nanjing, China, 2021. [CrossRef]
15. Li, J.; Wang, S.; Yu, J.; Wang, L.; Zhou, S. A modified CTAB protocol for plant DNA extraction. *Chin. Bull. Bot.* **2013**, *48*, 72–78. [CrossRef]
16. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef]
17. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amoes, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [CrossRef]
18. Rochette, N.C.; Catchen, J.M. Deriving genotypes from RAD-seq short-read data using Stacks. *Nat. Protoc.* **2017**, *12*, 2640–2659. [CrossRef] [PubMed]
19. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595. [CrossRef] [PubMed]
20. Hansen, M.H.; Lang, C.S.; Abildgaard, N.; Nyvold, C.G. Comparative evaluation of the heterozygous variant standard deviation as a quality measure for next-generation sequencing. *J. Biomed. Inform.* **2022**, *135*, 462–470. [CrossRef]

21. Wang, J.; Raskin, L.; Samuels, D.C.; Shyr, Y.; Guo, Y. Genome measures used for quality control are dependent on gene function and ancestry. *Bioinformatics* **2015**, *31*, 318–323. [CrossRef]
22. 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]
23. Hardy, G.H. Mendelian proportions in a mixed population. *Science* **1908**, *28*, 49–50. [CrossRef] [PubMed]
24. Nei, M. Genetic distance between populations. *Am. Nat.* **1972**, *106*, 283–292. [CrossRef]
25. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [CrossRef]
26. Geraldes, A.; Farzaneh, N.; Grassa, C.J.; Mckown, A.D.; Guy, R.D.; Mansfield, S.D.; Douglas, C.J.; Cronk, Q.C.B. Landscape genomics of *Populus trichocarpa*: The role of hybridization, limited gene flow, and natural selection in shaping patterns of population structure. *Evolution* **2014**, *68*, 3260–3280. [CrossRef]
27. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef]
28. 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]
29. Jakobsson, M.; Rosenberg, N.A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801–1806. [CrossRef] [PubMed]
30. Rosenberg, N.A. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* **2004**, *4*, 137–138. [CrossRef]
31. Mullenders, L.H.F. Solar UV damage to cellular DNA: From mechanisms to biological effects. *Photochem. Photobiol. Sci.* **2018**, *17*, 1842–1852. [CrossRef] [PubMed]
32. Morison, S.A.; Cramp, R.L.; Alton, L.A.; Franklin, C.E. Cooler temperatures slow the repair of DNA damage in tadpoles exposed to ultraviolet radiation: Implications for amphibian declines at high altitude. *Glob. Change Biol.* **2020**, *26*, 1225–1234. [CrossRef]
33. Martins, S.G.; Zilhão, R.; Thorsteinsdóttir, S.; Carlos, A.R. Linking Oxidative Stress and DNA Damage to Changes in the Expression of Extracellular Matrix Components. *Front. Genet.* **2021**, *12*, 673002. [CrossRef]
34. Dutta, S.; Mitra, M.; Agarwal, P.; Mahapatra, K.; De, S.; Sett, U.; Roy, S. Oxidative and genotoxic damages in plants in response to heavy metal stress and maintenance of genome stability. *Plant Signal. Behav.* **2018**, *13*, e1460048. [CrossRef]
35. Degtyareva, N.P.; Heyburn, L.; Sterling, J.; Resnick, M.A.; Gordenin, D.A.; Doetsch, P.W. Oxidative stress-induced mutagenesis in single-strand DNA occurs primarily at cytosines and is DNA polymerase zeta-dependent only for adenines and guanines. *Nucleic Acids Res.* **2013**, *41*, 8995–9005. [CrossRef]
36. Waneka, G.; Svendsen, J.M.; Havird, J.C.; Sloan, D.B. Mitochondrial mutations in *Caenorhabditis elegans* show signatures of oxidative damage and an AT-bias. *Genetics* **2021**, *219*, iyab116. [CrossRef]
37. Triest, L.; Rossum, F.V.; Sramkó, G.; Sierens, T.; Volkova, P. Over the hills and far away: Phylogeography and demographic migration history of a dispersal-restricted primrose (*Primula vulgaris*). *Front. Ecol. Evol.* **2024**, *12*, 1333726. [CrossRef]
38. Sun, L.; Geng, Q. Research methods and influencing factors of genetic diversity of rare and endangered plants. *Anhui Agric. Sci.* **2014**, *42*, 3793–3798. [CrossRef]
39. Zhang, D.; Jiang, X. Mating system evolution, resource allocation, and genetic diversity in plants. *Acta Phytoecol. Sin.* **2001**, *25*, 130–143.
40. Exposito-Alonso, M.; Booker, T.R.; Czech, L.; Gillespie, L.; Hateley, S.; Kyriazis, C.C.; Lang, P.L.M.; Leventhal, L.; Nogues-Bravo, D.; Pagowski, V.; et al. Genetic diversity loss in the Anthropocene. *Science* **2022**, *377*, 1431–1435. [CrossRef] [PubMed]
41. Romiguier, J.; Gayral, P.; Ballenghien, M.; Bernard, A.; Cahais, V.; Chenuil, A.; Chiari, Y.; Dernat, R.; Duret, L.; Faivre, N.; et al. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* **2014**, *515*, 261–263. [CrossRef]
42. Wittmann, M.J.; Gabriel, W.; Metzler, D. Genetic diversity in introduced populations with an Allee effect. *Genetics* **2014**, *198*, 299–310. [CrossRef]
43. Ellegren, H.; Galtier, N. Determinants of genetic diversity. *Nat. Rev. Genet.* **2016**, *17*, 422–433. [CrossRef]
44. Kling, M.M.; Ackerly, M.M. Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2017317118. [CrossRef]
45. Chen, Y.; Hou, Y.; Guo, Z.; Wang, W.; Zhong, C.; Zhou, R.; Shi, S. Applications of multiple nuclear genes to the molecular phylogeny, population genetics and hybrid identification in the mangrove genus *Rhizophora*. *PLoS ONE* **2015**, *10*, e0145058. [CrossRef] [PubMed]
46. Golenberg, E.M. Estimation of gene flow and genetic neighborhood size by indirect methods in a selfing annual, *Triticum dicoccoides*. *Evolution* **1987**, *41*, 1326–1334. [CrossRef] [PubMed]
47. Hamrick, J.L.; Godt, M.J.W.; Sherman-Broyles, S.L. *Factors Influencing Levels of Genetic Diversity in Woody Plant Species*; Springer: Dordrecht, The Netherlands, 1992. [CrossRef]
48. Demenocal, P.B.; Stringer, C. Human migration: Climate and the peopling of the world. *Nature* **2016**, *538*, 49–50. [CrossRef]

49. Stewart, J.R.; Stringer, C.B. Human evolution out of Africa: The role of refugia and climate change. *Science* **2012**, *335*, 1317–1321. [CrossRef]
50. Van Strien, M.J.; Holderegger, R.; Van Heck, H.J. Isolation-by-distance in landscapes: Considerations for landscape genetics. *Heredity* **2015**, *114*, 27–37. [CrossRef] [PubMed]
51. Sequeira, F.; Bessa-Silva, A.; Tarroso, P.; Sousa-Neves, T.; Vallinoto, M.; Goncalves, H.; Martínez-Solano, I. Discordant patterns of introgression across a narrow hybrid zone between two cryptic lineages of an Iberian endemic newt. *J. Evol. Biol.* **2020**, *33*, 202–216. [CrossRef] [PubMed]
52. Brousseau, L.; Foll, M.; Scotti-Saintagne, C.; Scotti, I. Neutral and adaptive drivers of microgeographic genetic divergence within continuous populations: The case of the neotropical tree Eperua falcata (Aubl.). *PLoS ONE* **2015**, *10*, e0121394. [CrossRef]
53. Pérez-Alquicira, J.; Wehncke, E.V.; García-Loza, G.A.; Carstens, B.C.; Domínguez, C.A.; Pérez-Ishiwara, R.; Molina-Freaner, F.E.; Zamora-Tavares, M.D.P.; Rico, Y.; Cabrera-Toledo, D.; et al. Geographic isolation and long-distance gene flow influence the genetic structure of the blue fan palm *Brahea armata* (*Arecaceae*). *J. Plant Res.* **2023**, *136*, 277–290. [CrossRef]
54. Yang, F.; Qin, A.; Li, Y.; Wang, X. Great genetic differentiation among populations of Meconopsis integrifolia and its implication for plant speciation in the Qinghai-Tibetan Plateau. *PLoS ONE* **2012**, *7*, e37196. [CrossRef] [PubMed]
55. Yardeni, G.; Tessler, N.; Imbert, E.; Sapir, Y. Reproductive isolation between populations of *Iris atropurpurea* is associated with ecological differentiation. *Ann. Bot.* **2016**, *118*, 971–982. [CrossRef]
56. Miles, L.S.; Rivkin, L.R.; Johnson, M.T.J.; Munshi-South, J.; Verrelli, B.C. Gene flow and genetic drift in urban environments. *Mol. Ecol.* **2019**, *28*, 4138–4151. [CrossRef]
57. Scotti, I.; Lalagüe, H.; Oddou-Muratorio, S.; Scotti-Saintagne, C.; Daniels, R.R.; Grivet, D.; Lefevre, F.; Cubry, P.; Fady, B.; González-Martínez, S.C.; et al. Common microgeographical selection patterns revealed in four European conifers. *Nat. Ecol. Evol.* **2023**, *32*, 393–411. [CrossRef]
58. Brown, N.L.; Peacock, M.M.; Ritchie, M.E. Genetic variation and population structure in a threatened species, the Utah prairie dog *Cynomys parvidens*: The use of genetic data to inform conservation actions. *Nat. Ecol.* **2016**, *6*, 426–446. [CrossRef]
59. Dutta, T.; Sharma, S.; Maldonado, J.E.; Panwar, H.S.; Seidensticker, J. Genetic Variation, Structure, and Gene Flow in a Sloth Bear (*Melursus ursinus*) Meta-Population in the Satpura-Maikal Landscape of Central India. *PLoS ONE* **2015**, *10*, e0123384. [CrossRef] [PubMed]
60. Ribeiro, F.E.; Baudouin, L.; Lebrun, P.; Chaves, L.J.; Brondani, C.; Zucchi, M.I.; Vencovsky, R. Population structures of Brazilian tall coconut (*Cocos nucifera* L.) by microsatellite markers. *Genet. Mol. Biol.* **2010**, *33*, 696–702. [CrossRef] [PubMed]
61. Des Roches, S.; Pendleton, L.H.; Shapiro, B.; Palkovacs, E.P. Conserving intraspecific variation for nature's contributions to people. *Nat. Ecol. Evol.* **2021**, *5*, 574–582. [CrossRef]
62. Martin, S.H.; Möst, M.; Palmer, W.J.; Salazar, C.; McMillan, W.O.; Jiggins, F.M.; Jiggins, C.D. Natural Selection and Genetic Diversity in the Butterfly *Heliconius melpomene*. *Genetics* **2016**, *203*, 525–541. [CrossRef] [PubMed]

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.