Genetic Diversity and Population Structure of *Dendrobium Nobile* in Southwest of China Based On Genotyping-by-Sequencing

Tao He (✉ hetao@cib.ac.cn)
CIB CAS: Chengdu Institute of Biology  [https://orcid.org/0000-0002-7342-9479](https://orcid.org/0000-0002-7342-9479)

Changrong Ye
Huazhi Biotech Co. Ltd

Qin Zeng
Huazhi Biotech Co. Ltd

Xiaoli Fan
CIB CAS: Chengdu Institute of Biology

Tianfang Huang
CIB CAS: Chengdu Institute of Biology

Research Article

Keywords: Dendrobium nobile Lindl., genetic diversity, population structure, genotyping-by-sequencing

Posted Date: November 30th, 2021

DOI: [https://doi.org/10.21203/rs.3.rs-1069533/v1](https://doi.org/10.21203/rs.3.rs-1069533/v1)

License: ☺ ① This work is licensed under a Creative Commons Attribution 4.0 International License.  [Read Full License](https://creativecommons.org/licenses/by/4.0/)

---

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.
Abstract

*Dendrobium nobile* Lindl. is one of the most important Orchid plants worldwide. The genotype-by-sequencing (GBS) method has now been widely used to access genetic diversity because of its high-throughput and cost-effective in molecular markers. The goal of this study was to employ the GBS technique for diversity evaluation of *D. nobile* and determine genetic differences between populations. A total of 129 accessions of *D. nobile* collected originally between 2019 and 2020 from 10 imitation-wild cultivated populations growing in Sichuan, Guizhou and Yunnan of southwestern China were sequenced, a total of 135G clean reads and a total of 836,786 SNPs of high quality data was yielded and were used for final analysis of genetic diversity and population structure. The quality value 20(Q20) ≥ 92.61%, the quality value 30(Q30) ≥ 82.38%. The GC contents distributed between 37.58% and 38.82%. It was also found that more transitions than transversions, and the ratio of transition/transversion varied from 1.804 to 1.911. By the methods of STRUCTURE, the most appropriate number was found to be k=3, all accessions of *D. nobile* were classified into three groups, excepts for 14 accessions belonging to admixed group. Phylogenetic tree and principal component analysis (PCA) were consistent with the result. The first two principal components explained a total of 23.25% of the variation by PCA. The genetic diversity of ML population showed the lower genetic diversity as indicated by the effective number of alleles (Ne) = 1.287, polymorphism information content (PIC) = 0.141, and Shannon’s information index (I) = 0.205, while WT population showed slightly higher genetic diversity by the Ne=1.512, PIC =0.256, and I =0.360. ML population and other nine populations (FB, FM, FX, LJ, SJ, SP, WL, WT and XM) were the most divergent between them respectively owing to all pairwise Fst values above 0.25, while FM population and FX population were considered identical because the pairwise Fst value was 0.0 between the two populations. Correlation analysis showed that highly significant correlation was observed between genetic distance and actual geographical distance (r = 0.854, P < 0.0001), indicating that the genetic differentiation of the 10 *D. nobile* populations conformed to the geographical isolation model. Analysis of molecular variance (AMOVA) revealed that the genetic variation was greater within populations (87.8%) than among populations (12.2%). This confirmed that intra-population variation was the main source of genetic variation in 10 *D. nobile* populations. The results also showed that Nm = 1.799 > 1, indicating that there was gene exchange between different populations. Analysis of unweighted pair-group method with arithmetic mean (UPGMA) suggested that the 10 populations were classified into three groups (Group I, Group II and Group III), Group III could be further divided into two subgroups (Group IIIa and Group IIIb). The results will not only provide valuable information for the level of genetic diversity of *D. nobile* growing in southwestern of China but also help for formulation of strategies for resource protection and utilization. Moreover, GBS appears as an efficient tool to detect intra-population variation.

Introduction

*Dendrobium nobile* Lindl., an epiphytic orchid, is found to be distributed in tropical and subtropical regions of Asia, including Thailand, Laos, Vietnam, Sikkim, Bhutan, Burma, and China. In China, it is found to be present in southern regions of the Tsinling Mountain, such as Yunnan, Guangxi, Guizhou, Sichuan, and Hubei provinces (Tsi et al. 1999). It is one of the most important and representative species in the genus *Dendrobium* with approximately 1500 species in the worldwide, and is considered as a favorite ornamental plants, known for unique type, bright color and long-lasting flowers. In addition, *D. nobile* plants has been used as precious and traditional Chinese medicine, which has the special efficacy of enhancing immunity, treating neurological diseases, digestive system diseases, ophthalmic diseases and anti-tumor owing to containing a variety of bioactive substances such as dendrobine, polysaccharide, flavonoids, and etc (Yang et al. 2014; Bhattacharyya et al. 2015). It is listed in Pharmacopoeia of the People’s Republic of China (Chinese pharmacopoeia Committee 2020). Therefore, *D. nobile* has high economic value, ecological value and scientific research value.

However, owing to the pressure of human activities, over-exploitation, natural habitat destruction and fragmentation, like other *Dendrobium* plants, *D. nobile* is facing the serious threat of endangered and is contained in Appendices I and II of Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) (Wang et al. 2018). In order to protect wild resources of *D. nobile*, the plants have been cultivated in China, particularly in southwestern regions of China for the market demand. Despite a long time of cultivation, few reports are yet available on genetic variation and population structure of *D. nobile*. Diversity in cultivated population is crucial for genetic improvement and for the related ornamental industry (Satya et al. 2015; Kang et al. 2015). It is necessary to study population structure and access the level of genetic diversity of the species. An analysis of genetic diversity using molecular marker is a trusted and reliable approach which could provide useful baseline information of various plants or population (Bhattacharyya et al. 2013).

Molecular markers have been developed and a variety of molecular markers have been used to detect genetic diversity and genetic relationships among species in the genus *Dendrobium* to date, including random-amplified polymorphic DNA (RAPD) (Zhu et al. 2011), inter simple sequence repeat (ISSR) (Yang et al. 2011), simple sequence repeat (SSR) (Lu et al. 2013; Kang et al. 2015), sequence
related amplified polymorphism (SRAP) (Feng et al. 2014), inter primer binding repeats site (iPBs) (Cui et al. 2020), target region amplification polymorphism (trap) (Feng et al. 2015), sequence-characterized amplified regions (scar) (Zheng et al. 2021), these markers are based on electrophoretic map. Other markers are based on sequence differences, including nuclear genome such as ITS, chloroplast genome, including matK, rbcl, psbA-trnH and mitochondrial genome (Ye et al. 2017; Wang et al. 2018; Nguyen et al. 2020). These studies showed that highly genetic differentiation and rich genetic diversity existed among different Dendrobium species. The revealed variabilities have allowed for better understanding of the extent of diversity present in the genus.

Efforts have also been made to assess the genetic diversity within species in the genus Dendrobium using different molecular markers. These studies revealed that genetic diversity and the variation within populations was greater than among populations in D. officinale (Jiang et al. 2016; Zhang et al. 2019; Xie et al. 2020) and in D. moniliforme (Ye et al. 2015). While most of the genetic variation found to be existed among populations in D. fimbriatum (Ma et al. 2009). Some studies aimed at understanding the genetic diversity of D. nobile using molecular markers for the analysis of population structure, and genetic variation had been reported in natural populations. The results showed that the genetic variation among natural populations from Northeast India was higher than within populations by Bhattacharyya et al. using SCoT (2013), RAPD (2015), ISSR and DAMD (2015) markers respectively. Other results also showed higher genetic differentiation among populations from Qinba mountain area in China using ISSR marker (Jiang et al. 2016). Whereas the genetic variation of D. nobile within populations from Yunnan Province, China was higher than among populations, using DALP marker (Zhang et al. 2013). In these studies, the natural population of D. nobile was mainly taken as the research object, and the sampling range were respectively located in Northeast India, Yunnan province or Qinba mountain area of China. However, the research on genetic variation of cultivated population was rarely reported.

In recent years, a helpful tool to enhance genetic analysis has emerged called genotyping-by-sequencing (GBS)(Chen et al. 2015), which is a new method based on next-generation-sequence (NGS) (Zhu et al. 2019) and massive sequencing which is able to discover a large number of single nucleotide polymorphisms (SNPs) (Wang et al. 2017). The main advantage of the method over previous approaches is that they can still work in the absence of a reference genome (Perea et al. 2016). It has dramatically reduced costs and time and provides a rapid, high-throughput and cost-effective tool for exploring plant genetic diversity on a genome-wide scale (Taranto et al. 2016). Therefore, GBS technology has been widely used for population structure and genetic diversity studies in various plant species (Taranto et al. 2016; Pavan et al. 2017; Wang et al. 2017; Cai et al. 2018; Guo et al. 2019; Morigengova et al. 2019; Niu et al. 2019; Wang et al. 2019; Zhou et al. 2019; Jia et al. 2020; Zhang et al. 2021).

To our knowledge, no population structure, and genetic diversity studies were available of D. nobile in previous study using GBS. In this study, we gathered 129 D. nobile accessions from 10 cultivated populations of Yunnan, Guizhou, and Sichuan provinces in southwestern China. The objectives of our study were to apply GBS technology (1) to discover SNPs for D. nobile species, (2) to assess the level of genetic diversity in a set of 129 accessions, (3) to evaluate the utility of the GBS in the genetic diversity analysis of D. nobile, an endangered medicinal plant.

**Materials And Methods**

**Plant Materials**

A total of 129 accessions of D. nobile were analyzed in this study. All samples were collected originally between 2019 and 2020 from 10 imitation-wild cultivated populations growing in Sichuan, Guizhou and Yunnan of Southwestern China. The species collected was identified by authors. The GPS coordinates of the 10 populations were recorded. Young leaves which were fresh, healthy and disease-free were cut off from randomly selected plants and immediately put into self-sealing bags containing dry Silica Gel, then they were taken back to the laboratory and stored in the refrigerator at -20°C prior to use. Provenances of the 10 populations used in this study were listed in Table 1.
Table 1
Provenances of the 10 populations used in this study

| Population code | Geographic location                                           | Longitude  | Latitude  | Altitude (m) | Sample Number |
|-----------------|---------------------------------------------------------------|------------|-----------|--------------|---------------|
| WL              | Wanglong Town in Chishui City of Guizhou Province             | E105°52´ | N28°30´   | 536           | 15            |
| FX              | Fuxing Town in Chishui City of Guizhou Province               | E105°45´ | N28°27´   | 388           | 15            |
| WT              | Wutong Town in Hejiang County of Sichuan Province             | E105°33´ | N28°30´   | 377           | 15            |
| FM              | Fengming Town in Hejiang County of Sichuan Province           | E105°52´ | N28°39´   | 430           | 14            |
| FB              | Fubao Town in Hejiang County of Sichuan Province              | E106°06´ | N28°44´   | 579           | 15            |
| SP              | Sipo Town in Yibin City of Sichuan Province                  | E104°26´ | N28°48´   | 407           | 11            |
| SJ              | Songjia Town in Yibin City of Sichuan Province                | E104°51´ | N28°44´   | 368           | 9             |
| LJ              | Liujiang Town in Hongya County of Sichuan Province            | E103°14´ | N29°43´   | 615           | 12            |
| XM              | Xiema Town in Jiajiang County of Sichuan Province             | E103°22´ | N29°46´   | 465           | 14            |
| ML              | Mengla Town in Mengla County of Yunnan Province               | E101°34´ | N21°29´   | 640           | 9             |

Genomic DNA Extraction

Genomic DNA was extracted from young leaf tissues of *D. nobile* using the cetyl-trimethyl-ammonium-bromide (CTAB) method or plant DNA Kit (Tiangen Biotech Co., Ltd., Beijing). The detection of DNA mainly included the 3 following kinds: (1) The integrity of DNA was verified by electrophoresis on 1% agarose gel. (2) The purity of DNA was measured by Nanodrop Spectrophotometer (nanodrop, np-1000 USA), the ratio of OD260/OD280 was between 1.8 and 2.0. (3) The concentration of DNA was quantified by the Qubit 2.0 fluorometer (Invitrogen, Carlsbad, USA).

Genotyping by Sequencing

Qualified DNA was first digested using the restriction enzymes (*Pst*I, *Msp*I), and barcoded adapters were ligated to the digested fragments, then PCR amplification was carried out, the PCR products were pooled, fragments between 350 and 600 bp were isolated from agarose gel and purified, the library was constructed. After the library was tested to be qualified, paired-end sequencing was performed on the Illumina novaseq6000 platform at Huazhi Biotech Co., Ltd. in Changsha, China.

All the raw pair-end sequencing data in FASTQ format were deposited into the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) with BioProject accession PRJNA747120. The original data were split, low-quality reads were removed, and clean reads were finally generated, which was a total of 135 G. The quality of all sequencing was evaluated using fastqc version 0.11.8 software. The results showed that all samples had no quality problems and the sequencing quality was good. 10000 reads were randomly selected from the FASTQ file of sequencing data and aligned to sequencing in NCBI data using blast version 2.2.26 software (Altschul et al. 1990) to detect the pollution of sequencing.

Using BWA version 0.7.13 software (Li and Durbin 2009), the clean reads were aligned to the reference genome sequence of *Dendrobium catenatum* Lindl. (GCF_001605985.2_ASM160598v2_genomic.fna) (Zhang et al. 2016).

Variation was detected using FreeBayes version 0.9.21 software (Ebbert et al. 2016) and filtered using bcftools software (parameters: view-m2-M2-v snps). SNP locus with allele frequency (MAF) less than 0.05 and missing data rate higher than 20% were filtered out using vcfhtools version 4.2 software with parameter: --minGQ 5 --maf 0.05 --max-missing 0.8 (Garrison and Marth 2012). SNPs information was statisticced using vcf-stats software (Danecek et al. 2011). After the filtering, 836,786 SNPs of 129 accessions were retained and used for further analysis.

Data Analysis

Based on the filtered SNPs, the population structure of 129 accessions was conducted using Admixture software, the number of subgroup varied from 1 to 12, the output of STRUCTURE was evaluated to determine the most appropriate k value, which will exhibit
the low cross-validation error compared to other k values. Samples were considered to belong to one of the subgroup when the value of the corresponding membership coefficient (Q) was higher than 0.6, if not, the sample was considered as admixed at assigned k.

The genetic distance matrix was calculated using plink software and phylogenetic tree was constructed using Neighbor Joining algorithm method. The entire accessions were divided into subgroups based on the major nodes of the phylogenetic tree.

Principal component analysis (PCA) was performed to assess genetic distinctness and redundancy, and to assess the sample association plots of the first two resulting principal components using GCTA version 1.26.0 software (Yang et al. 2011).

Genetic diversity parameters including observed heterozygosity (Ho), expected heterozygosity (He), effective number of alleles (Ne), polymorphism information content (PIC), Shannon’s information index (I), gene flow (Nm), were calculated using Arlequin version3.2 software (Excoffier et al. 2005). These parameters were used for the analysis and the comparison of the diversity of the 10 populations.

In addition, using Arlequin version3.2 software, genetic differentiation index (Fst) value was calculated, and the pairwise genetic distances among the 10 populations were evaluated, and the average values of Fst provided a measure of genetic distance between populations. In the meantime, unweighted pair-group method with arithmetic mean (UPGMA) dendrogram based on the pairwise genetic distances among the 10 populations obtained was used to assess inter-population differentiation and distinctness. An analysis of molecular variance (AMOVA) was performed to quantify the genetic variation present within and among the 10 populations.

The correlation coefficient between genetic distance and actual geographical distance was conducted using R version 3.6.3 software (Martin et al. 2014) to assess their correlation respectively.

**Results**

A total of 129 *D. nobile* accessions were sequenced using GBS technology in order to understand the genetic diversity and population structure of the 10 populations. After filtering and quality controlling, raw reads were split and finally generated a total of 135G clean reads. A total of 836,786 SNPs of high quality data was yielded and were used for final analysis. More than 60% of sequencing aligned to the reference genome, the depth of sequencing was 1×- 3×, and the coverage was basically more than 11%. Statistics on sequence data (Table 2) further showed that the number of total reads ranged from 6,497,132 (XM population) to 17,566,739 (WT population); The number of total bases ranged from 887,060,448 (XM population) to 2,373,845,107 (WT population); the quality value 20(Q20) ≥ 92.61%, with an average of 95.26%; the quality value 30(Q30) ≥ 82.38%, with an average of 87.60%; the GC contents distributed in a range of 37.58% (SJ population) to 38.82% (XM population). It was suggested that GBS was a valuable molecular method qualified for population genetic characterization of *D. nobile*.

| Populations | Total Reads | Total Bases | GC content (%) | Q20 (%) | Q30 (%) |
|-------------|-------------|-------------|----------------|---------|---------|
| FB          | 14,663,845  | 2,001,153,894| 38.33          | 92.61   | 82.38   |
| FM          | 14,258,947  | 1,928,206,553| 37.84          | 93.08   | 83.05   |
| FX          | 15,163,263  | 2,056,901,952| 38.06          | 96.26   | 89.76   |
| LJ          | 14,913,211  | 2,024,594,136| 38.27          | 94.29   | 85.47   |
| ML          | 12,749,467  | 1,723,895,648| 38.59          | 94.37   | 85.62   |
| SJ          | 15,522,683  | 2,102,755,840| 37.58          | 95.73   | 88.54   |
| SP          | 10,440,432  | 1,413,377,961| 38.23          | 96.51   | 90.26   |
| WL          | 16,553,946  | 2,249,397,485| 38.21          | 98.65   | 90.53   |
| WT          | 17,566,739  | 2,373,845,107| 38.00          | 96.34   | 89.78   |
| XM          | 6,497,132   | 887,060,448  | 38.82          | 96.24   | 89.63   |

It was shown in Table 3, from the genomic level, the total number of SNP sites in the 10 populations of *D. nobile* varied from 170,167 (XM population) to 255,053 (LJ population); The range of homozygous loci was 63,748 (SJ population)-141,624 (ML population), and
the percentage of that was between 28.54% (FB population) -75.03% (ML population); the range of heterozygous loci was 46,452 (ML population) -168,225 (FB population), and the percentage of that was 24.97% (ML population)-71.46% (FB population). It was also found that more transitions than transversions, and the ratio of transition/transversion varied from 1.804 to 1.911.

Table 3

| Populations | No. of SNP sites | No. of homozygous loci | No. of heterozygous loci | Homozygous Rate (%) | Heterozygous Rate (%) | No. of Transition | No. of Transversion | Ratio of Transition/Transversion |
|-------------|-----------------|------------------------|--------------------------|---------------------|-----------------------|-------------------|----------------------|-------------------------------|
| FB          | 234,918         | 66,692                 | 168,225                  | 28.54               | 71.46                 | 71,462            | 39,463               | 1.811                         |
| FM          | 234,451         | 80,601                 | 153,850                  | 34.20               | 65.80                 | 71,231            | 39,486               | 1.804                         |
| FX          | 233,722         | 76,868                 | 156,853                  | 32.70               | 67.30                 | 71,194            | 39,179               | 1.817                         |
| LJ          | 255,053         | 120,028                | 135,025                  | 47.12               | 52.88                 | 77,672            | 42,733               | 1.818                         |
| ML          | 188,076         | 141,624                | 46,452                   | 75.03               | 24.97                 | 57,950            | 30,329               | 1.911                         |
| SJ          | 211,806         | 63,748                 | 148,058                  | 30.26               | 69.74                 | 64,252            | 35,685               | 1.801                         |
| SP          | 188,446         | 66,465                 | 121,981                  | 36.40               | 63.60                 | 57,298            | 31,534               | 1.817                         |
| WL          | 230,915         | 76,260                 | 154,655                  | 32.73               | 67.27                 | 70,205            | 38,739               | 1.812                         |
| WT          | 241,389         | 100,930                | 140,459                  | 41.63               | 58.37                 | 73,330            | 40,537               | 1.809                         |
| XM          | 170,167         | 90,086                 | 80,081                   | 52.61               | 47.39                 | 52,102            | 28,183               | 1.849                         |

STRUCTURE analysis was used to study the population structure of *D. nobile* based on 836,786 high quality SNPs from the 129 accessions. STRUCTURE software was run and the cross validation error (CV error) was calculated with k ranging from 1 to 12 to determine the optimal k value and the number of groups (k) was plotted. The results showed that (Fig. 1): when the k value is from 1 to 2, the CV error value decreased rapidly, at k = 3, a distinct inflection point was observed, the CV error value decreased to the minimum value, and then gradually increased to 8, after a rapid, gentle and linear rising respectively, finally the CV error value reached the maximum value at k = 12. So the most appropriate number was found to be k=3, which produced the lowest CV error value compared to other k values.

As a consequence, a total of 129 *D. nobile* accessions were classified into three groups. To classify groups, accessions unequivocally were assigned to a group when the threshold of the membership coefficient (Q) was higher than 0.6, while those with the Q lower than 0.6 were assigned to the admixed group. It was shown in Fig. 2, the first group, highlighted in red, consisted of 11 accessions, accounting for 8.53% of all accessions, with 8 accessions from LJ population, 1 accession from FM:WT and XM population, respectively. The second group which was the largest group, highlighted in blue, accounting for 58.14% of all accessions, consisted of 75 accessions, including 15 accessions from FB population, 11 accessions from SP and WL population, respectively, 10 accessions from FX population, 9 accessions from FM and SJ population, respectively, 6 accessions from XM population and 4 from WT population. The third group, highlighted in green, consisted of 29 accessions, accounting for 22.48% of all accessions, including 9 accessions from ML population, 7 accessions from XM and WT population respectively, 3 accessions from WL population, and 1 accession from FM, FX and LJ population respectively. In addition, the 14 accessions, with Q value of each group was less than 0.6, accounting for 10.85% of all accessions, were in the admixed group, that could not be unequivocally assigned to any of the above three groups.

The genetic distance matrix of all *D. nobile* accessions from 10 populations was calculated by p-link software and cluster analysis was carried out to construct the phylogenetic tree, as shown in Fig. 3. The phylogenetic tree was further generated that classified 129 accessions into three major clusters, except for the remains of 14 accessions (gray in color) in admixed cluster also. By comparing STRUCTURE groups versus NJ-tree clusters, it was possible to observe that all the accessions in group I, II, III and admixed were distributed in the cluster I, II, III and admixed, respectively. Cluster I (red in color) was the smallest cluster consisting 11 accessions, most of them from LJ population, and others from FM:WT and XM population, respectively. Cluster II (blue in color) was the largest cluster consisting of 75 accessions from 8 populations respectively, except for ML and LJ populations. Cluster III (green in color) was
comparatively smaller cluster, consisting of 29 accessions from 7 populations, including ML, XM, WT, WL, FM, FX and LJ population respectively.

The clustering in phylogenetic tree revealed genetic relationship of 129 accessions from the 10 populations. The green cluster and the red cluster were close together, indicating that accessions in the two clusters had a close genetic relationship. The blue cluster was far away from the red and green cluster, indicating that the genetic relationship between them is also far away. Different color clusters had different degrees of accessions mixing of different populations, indicating that there was a certain phenomenon of gene exchange between different populations.

Further principal component analysis (PCA) was performed to identify patterns of genetic variation among individual accessions and the 10 populations. All accessions were divided into four groups in the PCA plot (Fig. 4a). “0(orange in color)”, “1(dark green in color)”, “2(cyan in color)” and “3(purple in color)” represented admixed group, group I, II, and III in the STRUCTURE group respectively. The first principal component explained 15.02% of variation and the second explained 8.23%, both explained 23.25% of the total variation. All accessions in group II and a small amount of accessions in group III were placed in the positive (upper) portion of the plot of PC2 axis, while all accessions in group I, in admixed group and most accessions in group III were placed in the negative (lower) portion of the plot of PC2 axis. Only group I in the positive (right) portion along the plot of PC1 axis, the other three groups were placed in the negative (left) portion along the plot of PC1 axis. Moreover, it had been found that accessions in the group III showed a loose distribution, while in the group I and II showed centralized in the PCA plot indicating the larger genetic diversity present in group III.

In Fig. 4b, an initial analysis of population structure was performed using PCA. The 10 populations were represented by 10 different colors in the plot. It had been found that accessions from LJ and WT population respectively showed high dispersal, indicating high and higher genetic diversity in the two populations; while accessions from ML population clustered tightly together, indicating lower genetic diversity and it could be distinguished from other populations at the PCA plot. Moreover, different populations, except for ML population, appeared to overlap and became undistinguishable clearly with PCA, indicating that their genetic background was similar.

From the above, we concluded that the results of STRUCTURE analysis, PCA, and the phylogenetic tree showed a high degree of similarity. All accessions could be classified into three groups and one admixed group and there was a consistent one-to-one match among the three methods.

In order to understand the genetic diversity of the ten populations, the diversity parameters were calculated in Table 4. It indicated that the mean values of $\bar{H}_o$, $\bar{H}_e$, $\bar{N}_e$, $\bar{P}IC$, and $\bar{I}$ ranged from 0.184 - 0.345, 0.315 - 0.351, 1.287 - 1.512, 0.141 - 0.256, 0.205 - 0.360, with an average of 0.276, 0.328, 1.441, 0.222, 0.284 respectively. The results suggested that the ML population showed the lower genetic diversity compared to other populations by the $\bar{N}_e$=1.287, $\bar{P}IC$=0.141, and $\bar{I}$=0.205, while the WT population showed the slightly higher genetic diversity by the $\bar{N}_e$=1.512, $\bar{P}IC$=0.256, and $\bar{I}$=0.360.
Table 4
Genetic diversity parameters for the 10 populations. Observed heterozygosity (Ho), expected heterozygosity (He), number of effective alleles (Ne), Polymorphic information content (PIC), Shannon diversity index (I)

| Populations | Ho   | He   | Ne   | PIC  | I    |
|-------------|------|------|------|------|------|
| FB          | 0.344| 0.339| 1.459| 0.232| 0.250|
| FM          | 0.273| 0.319| 1.494| 0.247| 0.332|
| FX          | 0.277| 0.315| 1.486| 0.242| 0.316|
| LJ          | 0.263| 0.327| 1.483| 0.238| 0.320|
| ML          | 0.184| 0.337| 1.287| 0.141| 0.205|
| SJ          | 0.345| 0.351| 1.395| 0.216| 0.210|
| SP          | 0.324| 0.342| 1.393| 0.198| 0.231|
| WL          | 0.2749| 0.320| 1.481| 0.243| 0.301|
| WT          | 0.2466| 0.320| 1.512| 0.256| 0.360|
| XM          | 0.2311| 0.315| 1.424| 0.209| 0.315|

The genetic difference among the 10 populations was estimated by comparative analysis of genetic differentiation coefficient (Fst) (Table 5). The pairwise Fst value and genetic distances among the 10 populations were different. The results revealed that the highest values of Fst were between ML population and other nine populations, with Fst values ranged from 0.271 to 0.446, respectively. Meantime, the Fst values between LJ population and FB, SJ, SP populations were 0.277, 0.283, 0.288 respectively, all above 0.25 indicating that the largest difference between them. A slightly higher pairwise Fst values (0.164, 0.161, 0.184, 0.198), suggested a high differentiation between LJ population and XM, FM, FX, WL populations. The pairwise Fst values between WT population and FM, FX, LJ, SJ, SP, WL, XM populations; XM population and FM, FX, SJ, SP, WL populations; FB population and WT, XM populations ranged from 0.061 to 0.147 respectively, suggested that there was a moderate amount of differentiation among these populations, indicating that these populations were relatively genetically distinct. The pairwise Fst values between other populations showed slight genetic differentiation, indicating these populations had lower variability. Particularly, FM and FX population were considered identical because the Fst value was 0.0 between the two populations.

Table 5
Pairwise values of genetic differentiation index (Fst) among the 10 populations of Dendrobium nobile Lindl.

| populations | FB | XM | FM | FX | LJ | SJ | SP | WL | WT | ML |
|-------------|----|----|----|----|----|----|----|----|----|----|
| FB          | 0.000|    |    |    |    |    |    |    |    |    |
| XM          | 0.141*| 0.000|    |    |    |    |    |    |    |    |
| FM          | 0.024*| 0.080*| 0.000|    |    |    |    |    |    |    |
| FX          | 0.016*| 0.091*| 0.000| 0.000|    |    |    |    |    |    |
| LJ          | 0.277*| 0.164*| 0.161*| 0.184*| 0.000|    |    |    |    |    |
| SJ          | 0.005*| 0.138*| 0.025*| 0.013*| 0.283*| 0.000|    |    |    |    |
| SP          | 0.012*| 0.134*| 0.031*| 0.019*| 0.288*| 0.003| 0.000|    |    |    |
| WL          | 0.023*| 0.091*| 0.009| 0.001| 0.198*| 0.019| 0.023| 0.000|    |    |
| WT          | 0.144*| 0.085*| 0.068*| 0.080*| 0.100*| 0.143*| 0.147*| 0.061*| 0.000|    |
| ML          | 0.413*| 0.334*| 0.336*| 0.356*| 0.309*| 0.446*| 0.446*| 0.356*| 0.271*| 0.000|

Note: Indicates significant pairwise Fst after sequential Bonferroni corrections (α = 0.05)

Furthermore, correlation between genetic distance and actual geographical distance of the 10 populations were also analyzed. Highly significant correlation was observed between genetic distance and actual geographical distance (r = 0.854, P < 0.0001), indicating that
the genetic differentiation of the 10 *D.nobile* populations conformed to the geographical isolation model.

In order to define the patterns of genetic variation and to estimate the variation components among the populations, an analysis of molecular variance (AMOVA) was performed based on the pairwise genetic distances between the 10 *D.nobile* populations. The AMOVA revealed that much greater variation within populations (28.96%+58.84%=87.8%) than among populations (12.2%, p<0.001) (Table 6). In other words, 28.96% and 58.84% of the variation was attributed to the differences among individuals within populations and within individuals respectively, whereas among the populations, account for 12.2% of the total variation. This confirmed that intra-population variation was the main source of genetic variation in the 10 *D. nobile* populations. The results also showed that *Nm* = 1.799 > 1, indicating that there was gene exchange between different populations.

Table 6
Analysis of molecular variance (AMOVA) of *Dendrobium nobile* Lindl. for the 10 populations

| Source of variation         | Degree of freedom | Sum of squares | Variance components | Percentage of variation | *P* value | Nm   |
|-----------------------------|-------------------|----------------|---------------------|-------------------------|-----------|------|
| Among populations           | 9                 | 4266042.36     | 14918.48            | 12.20                   | <0.001    | 1.799 |
| Among individuals within populations | 119              | 14577322.96    | 35409.85            | 28.96                   | <0.001    |      |
| Within individuals          | 129               | 8513143.50     | 71952.95            | 58.84                   | <0.001    |      |
| Total                       | 257               | 27356508.82    | 122281.28           |                         |           |      |

Analysis of unweighted pair-group method with arithmetic mean (UPGMA) dendrogram based on the AMOVA results revealed the 10 populations were classified into three distinct groups at the Phi statistic of 0.075 or more. Group I and Group II consisted of only one population of ML and LJ, respectively; Group III was the largest group, with 8 populations consisting of SJ, SP, FB, WL, FM, FX, XM, and WT populations. At the statistic of 0.05, Group III could be further divided into two subgroups named Group IIIa and Group IIIb. In Group IIIa, there were two populations, including XM and WT population; In Group IIIb, there were 6 populations, including SJ, SP, FB, WL, FM, and FX populations.

**Discussion**

Phenotypic traits and molecular markers are two different tools for assessing genetic diversity. In previous study, the diversity of 18 phenotypic traits of the same 10 populations was analyzed, and the results showed that inter-population variation was the main source of phenotypic traits variation of *D.nobile* (He et al. 2021). In the present study, the AMOVA analysis of the same 10 populations based on GBS molecular markers revealed that the intra-population variation was the major genetic variation of *D. nobile*, indicating morphological traits were often influenced by the environment.

At the same time, in previous study, the cluster analysis of the 10 populations of *D. nobile* based on 18 phenotypic traits showed that at the Euclidean distance of 6.5, the 10 populations could be divided into 3 groups, which generally showed the rule that populations with similar geographical location and latitude gathered together, and phenotypic traits showed the pattern of geographical variation. In this study, the same 10 populations of *D. nobile* were clustered by UPGMA method according to the genetic distance. At the genetic distance of 0.075, the 10 populations could also be divided into three groups. There were similarities and differences between clustering based on molecular marker and phenotypic traits. The similarities were as follows: in the two kinds of clusters, ML population, at the southernmost end of Yunnan Province was clustered alone. Differences were as follows: in phenotypic clustering, XM population and LJ population gathered together; while in molecular clustering, LJ population gathered solely, the XM population gathered with WT population firstly to form subgroup, and then with the other 6 populations clustering in another subgroup gathered to form a group, indicating that the genetic background of *D. nobile* resources from XM and WT population was similar, and accessions of XM population might be introduced from Wutong town in Hejiang county of Sichuan province. In addition, in phenotypic clustering, FX population was closer to WT population, but in molecular clustering, FX population was closer to FM population, indicating that the genetic background between the two populations was similar. The SP and SJ population clustered together in molecular clustering, which was grouped independently in phenotypic clustering. Similar reports had been reported that there were inconsistencies between molecular clustering and phenotypic clustering in *D. officinale* and other plants species (Zhang et al. 2019; Yang et al. 2016).

The results of cluster analysis were well supported by the analysis of genetic differentiation among the 10 populations based on the genetic differentiation coefficient (Fst). The pairwise values of Fst between ML population and the other 9 populations were the largest,
varied from 0.271 to 0.446, all above 0.25, indicating obvious genetic differences between them. It could explain why ML population is not clustered with the other 9 populations. The Fst value between FX population and FM population was 0, showing a high degree of genetic homogeneity, no genetic differentiation between the two populations.

In addition, the results were inconsistent with the report by Bhattacharyya et al. (2013, 2015), that a greater degree of variation existed among the populations compared to that within the populations of *D. nobile* in Northeast India using SCoT, ISSR, and DAMD markers. The other studies had reported the similar results of *D. nobile* (Zhang et al. 2013; Jiang et al. 2016). Similar reports in *D. moniliforme*, another species of the *Dendrobium* genus (Ye et al. 2017). The reasons were considered that there had been high genetic differentiation in the long-term evolution process owing to geographic isolation, fragmented distribution and lack of exchange between natural populations. In our study, the accessions of *D. nobile* were from cultivated populations, the *Nm* value for the 10 populations was 1.799, implicating that genetic exchange occurred among the *D. nobile* populations in the neighboring distributing regions.

A different level of genetic diversity of the 10 cultivated populations from southwestern China was observed by analysis of genetic parameters. For example, it was shown the higher genetic diversity in WT population, while the lower genetic diversity in ML population. Firstly, it could be related to its own breeding system which was asexual reproduction. In a separate population far away from other populations and within its own small population, the distribution area might have genetic drift due to geographical isolation and lack of gene exchange with other populations, thus reducing the level of genetic diversity of ML population. This result was consistent with the conclusion of previous report (Yan et al. 2015). Their study found that *D. nobile* from Hainan Island in China showed a low level of genetic diversity, and there was a high degree of genetic differentiation between *D. nobile* from mainland China. Due to geographical isolation, there were genetic differences and significant regional specificity. There was a relatively higher level of genetic diversity in WT population, the reason was that Wutong town located in Hejiang county of Sichuan province and at the junction with Chishui city of Guizhou province. It is convenient to communicate with Sichuan and Guizhou of two provinces so the genetic diversity in the population was high. Geographical factors also had an impact on genetic diversity. The 10 cultivated populations distributed in different regions, and the longitude and latitude span a large range. Different geographical locations also lead to different levels of genetic diversity to adapt to environment. The results were consistent with the reported that the latitude and altitude had a positive effect and average annual temperature had negative influence on the population genetic diversity of *Gymnocarpos przewalskii* (Xu et al. 2018). As Mousadik et al. (1996) reported that the genetic variations had been affected by its own breeding system, distribution area, habitat change, individual cell mutation, founder effect, species legacy areas produced in the Pleistocene, etc.

Molecular clustering still showed obvious geographical characteristics, indicating that genetic differentiation conformed to geographical isolation mode. Through correlation analysis, there was a significant correlation between genetic distance and geographical distance. The reason might be that the accessions collected in this study distributed widely, spanning 3 Provinces, including Sichuan, Guizhou and Yunnan province which distributed in southwestern China. It showed different climatic and ecological conditions in different geographical regions. There existed a different degree of genetic differentiation between these populations could be related to their adaptability to different environments. Similar reports had been reported in *D. officinale* (Yuan et al. 2013), as well as other families such as naked fruit wood (Xu et al. 2018) and cassava in Brazil (Carvalho and Schaal 2001).

In the current study, a combination of different three analysis methods was employed in studying the genetic diversity and population structure of *D. nobile* from the 10 populations in southwestern China. Interestingly, the three methods, including STRUCTURE analysis, phylogenetic tree analysis, and principal component analysis (PCA) in general showed the same results and supported each other. The all accessions of *D. nobile* could be divided into three groups using the three methods respectively, except for the 14 accessions which were in admixed group. Moreover, the accessions classified by the three methods correspond to each other one by one. In other words, the three methods play a mutually reinforcing role.

In this study, comprehensive analysis showed that GBS technique could be suitable for the genetic diversity and population structure studies of *D. nobile*, an endangered medicinal plant. In the future, with the continuous improvement of the method, it would be optimized and used in the *D. nobile* for various purposes, such as construction of high density genetic maps, genetic mapping of complex traits through Genome-Wide Association Studies (GWAS), and estimation of breeding values in genetic selection as reported by Taranto et al. (2016).

**Declarations**

**Author contributions**

Page 10/16
Tao He contributed to experimental design, data analysis, and writing of this article. Changrong Ye and Qin Zeng contributed to GBS sequencing and bioinformatics. Xiaoli Fan and Tianfang Huang contributed to participating in experiments, and data analysis.

**Funding**

This research was funded by National Natural Science Foundation of China project, number: 31870308, and Science and technology poverty alleviation of Chinese Academy of Sciences project, number: KFJ-FP-202006.

**Conflict of interest**

The authors declare that they have no conflicts of interest or competing interests.

**Consent for publication**

All authors consent to publish.

**References**

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
2. Bhattacharyya P, Kumaria S, Tandon P (2013) Start codon targeted (SCoT) marker reveals genetic diversity of *Dendrobium nobile* Lindl., an endangered medicinal orchid species. Gene 529:21–26
3. Bhattacharyya P, Kumaria S (2015) Molecular characterization of *Dendrobium nobile* Lindl., an endangered medicinal orchid, based on randomly amplified polymorphic DNA. Plant Syst Evol 301:201–210
4. Bhattacharyya P, Kumaria S, Tandon P (2015) Applicability of ISSR and DAMD markers for phyto-molecular characterization and association with some important biochemical traits of *Dendrobium nobile*, an endangered medicinal orchid. Phytochemistry 117:306–316
5. Cai L, Yang H, Wang Y, Li TX, Chen GD (2018) Analysis of genetic diversity of tobacco germplasm resources based on SNP markers via genotyping-by-sequencing technology. Chinese Tobacco Science 39(5):17–24
6. Carvalho LJCB, Schaal BA (2001) Assessing genetic diversity in the cassava (*Manihot esculenta* Crantz) germplasm collection in Brazil using PCR-based markers. Euphytica 120:133–142
7. Chen TQ, Huang F, Li WZ, Jiang XL, Wang W, Liu DC, Yang WL, Zhang AM, Zhang LY (2015) Investigation and analysis of genetic diversity and population structure for wheat germplasms resistant to stripe rust in southwest China. Journal of Plant Genetic resources 16(6):1157–1167
8. Chinese Pharmacopoeia Commission (2020) Pharmacopoeia of the People's Republic of China, vol I. China Medical Science and Technology Press, Beijing pp
9. Cui XQ, Tang X, Huang CY, Deng JL, Li XL, Lu JS, Zhang ZB (2021) Genetic diversity analysis and fingerprinting construction of *Dendrobium* germplasm resources by iPBS marker. Chinese Journal of Tropical Crops 42(2):317–324
10. Danecek P, Auton A, Abecasis G et al (2011) The variant call format and VCFtools. Bioinformatics 27(15):2156–2158
11. Ebbert MTW, Wadsworth ME, Staley LA et al (2016) Evaluating the necessity of PCR duplicate removal from next-generation sequencing data and a comparison of approaches. BMC Bioinformatics 17(Suppl 7:239):492–543
12. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1:47–50
13. Feng SG, Lu JJ, Gao L, Liu JJ, Wang HZ (2014) Molecular phylogeny analysis and species identification of *Dendrobium* (Orchidaceae) in China. Biochem Genet 52:127–136
14. Feng SG, He RF, Yang S, Chen Z, Jiang MY, Lu JJ, Wang HZ (2015) Start codon targeted (SCoT) and target region amplification polymorphism (TRAP) for evaluating the genetic relationship of *Dendrobium* species. Gene 567:182–188
15. Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. Quantitative Biology 20:1–9
16. Guo Y, Qiao DH, Yang C, Li Y, Chen ZW, Chen J (2019) Genetic diversity of old tea plant resources in Jiuan city of Guizhou province, using genome-wide SNP. Journal of Plant Genetic Resources 20(1):26–36
17. He T, Fan XL, Lu L, Huang TF (2021) Phenotypic trait diversity of *Dendrobium nobile* from different provenances and their correlations with geographical factors. Journal of Plant Resources and Environment 30(2):1–11. DOI
18. Jia CR, Zhu XL, Wang L, Zhou CP, Weng QJ, Gan SM, Li FG (2020) SNPs discovery and phylogenetic analysis in Eucalypts using ddGBS. Journal of Central South University of Forestry & Technology 40(10):101–108
19. Jiang AM, Cao J, Cai GL (2016) Genetic relationship of Dendrobium and genetic diversity of Dendrobium nobile in Qinling-Daba mountains revealed by ISSR. Acta Bot Boreal-Occident Sin 36(10):1977–1983
20. Jiang JL, Ye W, Li YQ, Fan ZM, Lei FG, Zhang R, Zhou JJ (2016) Genetic relationships among wild Dendrobium officinale populations in Taining, Fujian by ISSR. Journal of Tropical and Subtropical Botany 24(3):259–266
21. Kang JY, Lu JJ, Qiu S, Chen Z, Liu JJ, Wang HZ (2015) Dendrobium SSR markers play a good role in genetic diversity and phylogenetic analysis of Orchidaceae species. Sci Hort 183:160–166
22. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheelers transform. Bioinformatics 25(14):1754–1760
23. Lu JJ, Kang JY, Feng SG, Zhao HY, Liu JJ, Wang HZ (2013) Transferability of SSR markers derived from Dendrobium nobile expressed sequence tags (ESTs) and their utilization in Dendrobium phylogeny analysis. Sci Hort 158:8–15
24. Ma JM, Yin SH (2009) Genetic diversity of Dendrobium fimbriatum (Orchidaceae), an endangered species, detected by inter-simple sequence repeat (ISSR). Acta Botanica Yunnanica 31(1):35–41. DOI
25. Martin SH, Davey JW, Jiggins CD (2014) Evaluating the use of ABBA–BABA statistics to locate introgressed loci. Mol Biol Evol 32(1):244–257
26. Mousadik AE, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (Argania spinose (L.) Skeel] endemic to Morocco. Theor Appl Genet 92:832–839
27. Morigengaowa, Shang H, Liu BD, Kang M, Yan YH (2019) One or more species? GBS sequencing and morphological traits evidence reveal species diversification of Sphaeropteris brunoniana in China. Biodiversity Science 27(11):1196–1204
28. Nguyen N-H, Vu H-T, Le N-D, Nguyen T-D, Duong H-X, Tran H-D (2020) Molecular identification and evaluation of the genetic diversity of Dendrobium species collected in Southern Vietnam. Biology 9:76
29. Niu SZ, Song QF, Koiwa H, Qiao DH, Zhao DG, Chen ZW, Liu X, Wen XP (2019) Genetic diversity, linkage disequilibrium, and population structure analysis of the tea plant (Camellia sinensis) from an origin center, Guizhou plateau, using genome-wide SNPs developed by genotyping-by-sequencing. BMC Plant Biol 19:328
30. Pavan S, Marcotrigiano AR, Ciani E, Mazzeo R, Zonno V, Ruggieri V, Lotti C, Ricciardi L (2017) Genotyping-by-sequencing of a melon (Cucumis melo L.) germplasm collection from a secondary center of diversity highlights patterns of genetic variation and genomic features of different gene pools. BMC Genomics 18:59
31. Perea C, Hoz JFDL, Cruz DF, Lobaton JD, Izquierdo P, Quintero JC, Raatz B, Duitama J (2016) Bioinformatic analysis of genotype by sequencing (GBS) data with LNSEP. BMC Genomics 17(Suppl 5, 498):539–551
32. Taranto F, D’Agostino N, Greco B, Cardi T, Tripodi P (2016) Genome-wide SNP discovery and population structure analysis in pepper (Capsicum annuum) using genotyping-by-sequencing. BMC Genomics 17:943
33. Satya P, Karan M, Jana S, Mitra S, Sharma A, Karmakar PG, Ray DP (2015) Start codon targeted (SCoT) polymorphism reveals genetic diversity in wild and domesticated populations of ramie (Boehmeria nivea L. Gaudich), a premium textile fiber producing species. Meta Gene 3:62–70
34. Tsi ZH, Chen SC, Luo YB, Zhu GH (1999) Orchidaceae (3). In: Tsi ZH (ed) Angiospermae, Monotyledoneae, Flora Reipublicae Popularis Sinicae. Science Press, Beijing
35. Wang SS, Hou FX, Zhao JR, Cao J, Peng C, Wan DG, Guo JL (2018) Authentication of Chinese herbal medicines Dendrobium species and phylogenetic study based on nrDNA ITS sequence. Int J Agric Biol 20:369–374
36. Wang X, Gao M, Wu LW, Wang YD, Liu HC, Chen YC (2019) A study on population genetic diversity among Litsea populifolia (Hems.) gamble in mount Emei area. Journal of Plant Genetic Resources 20(2):359–369
37. Wang XK, Jiang D, Sun ZZ (2017) Study on phylogeny of 240 Mandarin accessions with genotyping-by-sequencing technology. Scientia Agricultura Sinica 50(9):1666–1673
38. Xie JK, Zuo JH, Huang YH, Li CS, Chen YL (2020) The origin and germplasm collection for cultivated Dendrobium officinale K. Kimura & Migo individuals revealed by EST-SSR markers. Genet Resour Crop Evol 67:1209–1219
39. Xu ZP, Wan T, Cai P, Zhang XM, Yi WD, Wan YC (2018) Study on the relevance between genetic diversity of Gymnocarpos przewalskii populations and geographic climatic factors. ACTA AGRESTIA SINICA 26(1):70–76
40. Yang CY, Li XL, Wang YQ, Tang DY, Zhang ZL, Gao WW (2011) ISSR analysis of cultivated medicinal species of Dendrobium Sw. Chin Agric Sci Bull 27(04):148–152
41. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88:76–82
42. Yang X, Liu F, Zhang Y, Cheng YF, Xue LB, Chen XH (2016) Study on the genetic diversity of eggplant germplasm with SSR markers. Genomics and Applied Biology 35(12):3450–3457
43. Yang S, Gong QH, Wu Q, Li F, Lu YF, Shi JS (2014) Alkaloids enriched extract from Dendrobium nobile Lindl. attenuates tau protein hyperphosphorylation and apoptosis induced by lipopolysaccharide in rat brain. Phytomedicine 21:712–716
44. Ye MR, Hou BW, Luo J, Yan WJ, Liu W, Ding XY (2015) Genetic diversity and conservation of the endangered herb Dendrobium moniliforme based on amplified fragment length polymorphism markers. Sci Hort 189:51–58
45. Ye MR, Liu W, Xue QY, Yan WJ, Hou BW, Luo J, Ding XY (2017) Phylogeography of endangered Dendrobium moniliforme in East Asia based on mitochondrial DNA sequence variations. Biodivers Conserv 26:1659–1674
46. Zhang MY, Yu H, Yuan F (2013) Study on population genetic variation of Dendrobium nobile in Yunnan by DALP. China J Chin Materia Med 38(22):3866–3870. DOI:
47. Zhang GQ, Xu Q, Bian C et al (2016) The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. Sci Rep 6:19029
48. Zhang ZY, Zhou ML, Liang JP, Huang PP, Zhang C (2019) Genetic diversity analysis of Dendrobium officinale germplasm resources in Guanzhaishan. Molecular Plant Breeding 17(18):6179–6185
49. Zhang X, Zhang XJ, Ma YP, Li ZH, Wan YM, Ma H (2021) Genetic diversity assessment of Rhododendron sinofalconeri with genotyping by sequencing (GBS). Bulletin of Botanical Research 41(3):429–435. DOI
50. Zheng KX, Cai YC, Chen WJ, Gao YD, Jin JJ, Wang HZ, Feng SG, Lu JJ (2021) Development, identification, and application of a germplasm specific SCAR Marker for Dendrobium officinale Kimura et Migo. Front Plant Sci 12:1–9
51. Zhou PP, Yan HH, Peng YY (2019) Hexaploid ancestor of cultivated hexaploid oats inferred from high throughput GBS-SNP markers. ACTA AGRONOMICA SINICA 45(10):1604–1612
52. Zhu SN, Zhou ZH, Feng SG, Wang S, Suo NN, Wang HZ (2011) Analysis on the genetic diversity among 31 species of Dendrobium based on RAPD marker. J of Hangzhou Normal University (Natural Science ed.) 10(4):333–339
53. Zhu SL, Niu EL, Shi AN, Mou BQ (2019) Genetic diversity analysis of olive germplasm (Olea europaea L.) with genotyping-by-sequencing technology. Front Genet 10:1–11

Figures
Figure 1

A curve graph of cross-validation (CV) error values at different k value

Figure 2

A structure plot of the 129 accessions of Dendrobium nobile Lindl. by k=3 using 836,786 SNPs
Figure 3

Phylogenetic dendrogram of 129 accessions of Dendrobium nobile Lindl. from the ten populations based on 836,786 SNPs
Figure 4

Principal component analysis (PCA) plots of the 129 accessions Dendrobium nobile Lindl. based on 836,786 SNPs a Numbers (0-3) represent subgroup derived from structure analysis, colors represent different subgroup b Colors represent the 10 populations used in the study

Figure 5

UPGMA dendrogram of the 10 populations based on pairwise Fst value