

Unveiling genetic differentiation and admixture in modern *Hedychium* germplasm: a population genomics perspective from RAD-seq

Xiaoting Luo¹, Ping Bao^{1,2}, Changfu Cai¹, Mengyuan Wang¹, Yuxuan Xie^{1,2}, Jinying Wang¹ and Bangping Cai^{1,2*}

¹ Fujian Institute of Subtropical Botany (Xiamen Botanical Garden), Xiamen 361003, China

² College of Landscape Architecture and Art, Fujian Agriculture and Forestry University, Fuzhou 350002, China

* Correspondence: caibangping26@fjib.com (Cai B)

Abstract

The genus *Hedychium* has high ornamental and economic value, but its morphological classification remains contentious and the genetic backgrounds of many cultivars are unclear. In this study, we initially sequenced 168 *Hedychium* samples representing various species accessions, cultivars, artificial hybrids, and accessions of uncertain identity from China, the USA, and Thailand, using restriction site-associated DNA sequencing (RAD-seq). One low-quality sample was excluded from the principal component analysis (PCA), admixture, and D-statistic analyses, and its terminal branch was pruned from the displayed neighbor-joining (NJ) topology. In total, 856.92 Gb of high-quality data were generated, and 647,693 high-quality single-nucleotide polymorphisms (SNPs) were used for constructing the NJ tree, 265,365 linkage disequilibrium (LD)-pruned SNPs were used for PCA and admixture analyses, and 100,000 randomly sampled genome-wide SNPs were used for D-statistic analysis. The NJ tree resolved eight major clades, which were broadly supported by PCA and admixture analyses. D-statistic analysis further detected significant asymmetric allele sharing among selected accessions. Cultivated materials were associated with multiple genetic backgrounds represented by the sampled species accessions, including groups related to *H. coronarium*, *H. coccineum*, *H. gardnerianum*, and *H. forrestii*. Artificial hybrid accessions generally displayed mixed genetic components consistent with their recorded hybrid backgrounds. Several accessions of uncertain identity were assigned to closely related genetic groups, providing clues for their further identification. Our study provides a genome-scale population genetic framework for understanding genetic differentiation, admixture patterns, and genetic affinities within major *Hedychium* germplasm, offering molecular insights for the classification, conservation, and future breeding.

Keywords: *Hedychium*, RAD-seq, SNP, Genetic diversity and structure

Citation: Luo X, Bao P, Cai C, Wang M, Xie Y, et al. 2026. Unveiling genetic differentiation and admixture in modern *Hedychium* germplasm: a population genomics perspective from RAD-seq. *Ornamental Plant Research* 6: e035 <https://doi.org/10.48130/opr-0026-0026>

Introduction

The genus *Hedychium* (Zingiberaceae) comprises approximately 80 species of perennial herbs, with a major center of diversity in the Himalayan region and a natural distribution extending across tropical and subtropical Asia^[1,2]. Valued for their striking floral morphology, rich fragrance, and diverse biological activities, *Hedychium* species have been widely introduced and cultivated worldwide, with over 110 cultivars developed for ornamental, medicinal, and aromatic applications^[3]. However, taxonomic delineation of the genus based on traditional morphology remains contentious, largely because of ambiguous interspecific morphological boundaries and frequent intra- and interspecific hybridization. These complexities have resulted in unclear species limits and uncertain genetic origins for many cultivars^[4–6]. Previous molecular studies using markers such as isozymes, random amplified polymorphic DNA (RAPD), inter-simple sequence repeat (ISSR), and simple sequence repeats (SSRs) have provided preliminary insights into the genetic diversity of *Hedychium*^[7–12]. Nevertheless, the low resolution and limited genome-wide coverage of these conventional markers have constrained a deeper understanding of the population's genetic structure and complex breeding history, particularly within the cultivated gene pool. Consequently, several fundamental questions remain unresolved: (1) Which sampled *Hedychium* species show close genetic affinities with the major groups of modern *Hedychium* cultivars; (2) how admixture patterns potentially associated

with historical hybridization are reflected in the contemporary genetic architecture of modern *Hedychium* cultivars; and (3) whether high-density genomic markers can resolve genetic relationships among major germplasm resources and clarify the placement of unidentified accessions.

Restriction site-associated DNA sequencing (RAD-seq) has emerged as an efficient approach for developing genome-wide molecular markers, offering high throughput, accuracy, and rapid turnaround^[13,14]. By generating high-density single-nucleotide polymorphism (SNP) datasets, RAD-seq enables robust analyses of population structure, phylogenetic relationships, and genetic admixture, making it particularly suitable for elucidating complex breeding histories^[15]. This method has also been widely used to investigate admixture, introgression, and allele sharing patterns in nonmodel plant groups^[16,17]. Despite its advantages, a comprehensive RAD-seq-based population genomic study examining genetic structure and genetic affinities between species accessions and cultivated *Hedychium* germplasm is still lacking.

In this study, we initially collected 168 individual samples representing 35 *Hedychium* germplasm accessions and applied RAD-seq to generate a comprehensive set of polymorphic SNP loci. Using these high-quality SNPs, we conducted phylogenetic reconstruction, principal component analysis (PCA), population structure inference, and D-statistics analysis to elucidate the genetic relationships, population subdivision, and asymmetric allele sharing among *Hedychium* accessions. Our objectives were to identify *Hedychium* species

showing close genetic affinities with the major cultivated groups, characterize admixture patterns in modern cultivars and assess the relationships of these patterns with known or possible hybrid origins, and clarify the genetic placement of major and unidentified germplasm resources. The findings from this study will provide a crucial molecular foundation for the taxonomic clarification, breeding, conservation, and utilization of *Hedychium* germplasm.

Materials and methods

Plant materials

In total, 35 *Hedychium* germplasm accessions were selected for this study, comprising eight species accessions, eighteen cultivars, five artificial hybrids, and four accessions of unknown identity. These accessions, originating from China, the USA and Thailand, are maintained in the Zingiberales Garden at the Xiamen Botanical Garden (Supplementary Table S1). Accessions with uncertain identities, including Unknown 1–4 and provisional *Hedychium* sp. materials, were retained in the analysis to evaluate their genetic placement within the broader germplasm collection; their available background information is provided in Supplementary Table S1. From these accessions, 168 individual plant samples were initially collected, with six individual plants sampled for most accessions and one to five for those with limited availability. The 'Orange Brush' cultivar was an exception, with 11 individual plants collected. Approximately 0.5 g of healthy young leaf tissue was sampled from each plant, rapidly frozen in liquid nitrogen, and stored at -80°C for subsequent DNA extraction.

Genomic DNA extraction and quality assessment

Total genomic DNA was isolated from the *Hedychium* leaf samples using a plant DNA extraction kit (DNA BR). The integrity of the extracted DNA was verified by electrophoresis on a 1% agarose gel. DNA concentration was measured using either a Qubit Fluorometer or a microplate reader. Only DNA samples meeting quality standards were used for subsequent library construction.

Library construction and sequencing

Library preparation and sequencing were performed by BGI-Shenzhen (Shenzhen, China). The specific procedure was as follows: Genomic DNA (0.1–1 μg in a 50- μL reaction volume) was digested with the EcoRI enzyme (20 U per reaction) at 37°C for 15 min, followed by ligation to a Solexa P1 adapter containing a sample-specific barcode. The barcoded samples were then pooled and physically sheared to fragment sizes ranging from 300 to 500 bp. Target fragments within this size range were recovered by excision from a 1% agarose gel. The recovered fragments underwent end repair, A-tailing, and ligation to a Y-shaped P2 adapter. A 5- μL aliquot of the ligation product was used for polymerase chain reaction (PCR) amplification. Finally, the library's fragments, sized between 350 and 550 bp, were purified via gel electrophoresis^[18]. Library quality was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and quantitative PCR (qPCR). Libraries passing quality control were subjected to sequencing.

Raw data processing and quality control

Raw sequencing data underwent stringent quality control to obtain high-quality clean data for downstream analysis. Reads were first demultiplexed and assigned to their respective samples

according to their unique 5' barcodes (4–8 bp in length). The barcode sequences were then trimmed, and the presence of the correct EcoRI cut-site remnant (AATTC) was verified; only reads containing the correct site were retained. Further filtering removed adapter-contaminated reads, reads in which more than 40% of bases had a Phred quality score of ≤ 15 , and reads containing more than 10% ambiguous bases (N)^[19].

SNP calling and preprocessing

The demultiplexing and quality-control procedures described above were implemented using the process_radtags module of the STACKS pipeline (v1.44)^[20]. Subsequently, the ustacks module was used to assemble loci within each sample, followed by cstacks to build a catalog of loci across all samples. The sstacks module was then used to match loci from each sample against the catalog. Finally, genotyping was performed using the populations module, yielding a dataset of 647,693 high-quality SNPs. This dataset was used for constructing the neighbor-joining (NJ) tree.

To address potential biases from linked loci in analyses that assume interlocus independence, a separate SNP dataset was generated for PCA and admixture analyses. Starting from the original STACKS variant call format (VCF) output, loci were filtered using a site-level genotype missingness threshold of $< 50\%$, a minimum genotype depth of $5\times$, and a requirement of at least two observed alleles, yielding 1,441,743 SNPs. Linkage disequilibrium (LD) pruning was then performed using PLINK v2.0 with a sliding window of 50 SNPs, a step size of 10 SNPs, and an r^2 threshold of 0.2, resulting in a final set of 265,365 LD-pruned SNPs for PCA and admixture analyses^[21]. The original VCF was used as the source dataset for D-statistic analysis. Because Sample 30-6 yielded only 0.07 Gb of clean data, it was removed from both the LD-pruned dataset used for PCA and admixture analyses and the VCF used as input for D-statistic analysis. Both updated datasets therefore contained 167 samples.

Population genetic structure analysis

The corresponding filtered SNP datasets were used for population genetic analyses as described below. An NJ tree was constructed from the 647,693 high-quality SNP dataset using TreeBest (<https://treesoft.sourceforge.net/treebest.shtml>), based on a pairwise genetic distance matrix calculated with the p-distance method, to visualize the genetic similarity and clustering patterns among accessions. Topology robustness was assessed with 1,000 bootstrap replicates. Because RAD-seq data from a closely related external genus were not available in this study, *H. villosum* was used as a functional outgroup to root the NJ tree. This choice was based on a previous molecular phylogenetic study of Thai *Hedychium*, in which *H. villosum* was separated from the group containing *H. coronarium*, *H. flavescens*, *H. coccineum*, and *H. forrestii*^[9]. The NJ tree was visualized and annotated using iTOL (<https://itol.embl.de/>)^[22], with Sample 30-6 removed from the final visualization because it had insufficient sequencing data.

PCA was performed using PLINK v2.0 on the 265,365 SNPs after LD pruning. Because *H. villosum* (the designated outgroup) and Unknown 3 were extreme outliers that compressed the spatial distribution of the remaining samples in the preliminary analysis (Supplementary Fig. S1), these two accessions were excluded, and the final PCA was conducted on the remaining 165 samples. The PCA results were visualized using the ggplot2 package in R v4.4.1^[23].

Ancestral genetic components and population structure were inferred using ADMIXTURE software on the 167 sample LD-pruned SNP dataset containing 265,365 SNPs^[24]. Cross-validation (CV) error

was calculated for K values from 2 to 20 to evaluate the model's fit. Because the lowest CV error at high K values may reflect fine-scale differentiation among closely related cultivars or repeated individuals rather than the major genetic groupings, the main value for biological interpretation was determined by integrating the CV error trends with the concordance of inferred components with the groupings identified by the NJ tree and PCA. The results were visualized using TBtools (version 2.357)^[25].

D-statistic analysis

D-statistic analysis was performed using Dsuite (v0.5) to test for asymmetric allele sharing among *Hedychium* accessions under the ABBA-BABA framework^[26–28]. The test was based on the four-taxon model ((P1, P2), P3), outgroup), in which significant deviations from equal ABBA and BABA site patterns indicate allele sharing inconsistent with a strictly bifurcating tree. Such signals may reflect historical introgression, incomplete lineage sorting, or other sources of discordance among taxa. Therefore, the D-statistic was used here as a complementary test for asymmetric allele sharing rather than as a direct estimator of the parental contribution or the hybridization direction.

To balance computational efficiency with statistical power, 100,000 biallelic SNP sites were randomly sampled from the genome-wide SNP set using BCFtools (v1.9). The DtriosParallel module in Dsuite was then used to calculate genome-wide D-statistics, Z-scores, and *p*-values for all possible ingroup trios (P1, P2, P3), with *H. villosum* used as the outgroup. *H. villosum* was selected because it was designated as the functional outgroup in the NJ tree analysis, whereas Unknown 3 was not used as an outgroup because of its uncertain identity and highly divergent genetic position. Test combinations involving Unknown 3 were excluded from downstream interpretation because they produced unstable statistical estimates. Results with $|Z| > 3$ were considered to be significant and were further interpreted as complementary evidence for asymmetric allele sharing among accessions.

Results

Quality control and filtration of raw data

Following demultiplexing and stringent quality control, 856.92 Gb of high-quality clean data was obtained in total from the raw sequencing data. The average output per sample was 5.1 Gb of high-quality sequence data. The average percentage of bases with a Phred quality score ≥ 30 (Q30) and the average guanine-cytosine (GC) content across all samples were 93.13% and 38.12%, respectively (Supplementary Table S2), indicating high sequencing quality suitable for subsequent bioinformatic analyses. Notably, Sample 30-6 yielded substantially less data (0.07 Gb) than other samples, suggesting a potential issue during library construction or sequencing. Because of this insufficient sequencing output, Sample 30-6 was excluded from the PCA, admixture, and D-statistic datasets and was removed from the final NJ tree visualization.

SNP detection results

In total, 42,212,223 raw SNPs were identified across the 168 *Hedychium* samples. The number of SNPs detected per sample varied widely, ranging from 4,445 (30-6) to 510,277 (4-6), with an average of 251,263 SNPs per sample. The homozygous rate across all samples ranged from 9.10% to 91.90%, with an overall average of 57.63%. Conversely, the heterozygous rate ranged from 8.10% to

90.90%, with an average of 42.37% (Supplementary Table S3). The STACKS pipeline generated a final set of 647,693 high-quality SNP loci, which was used for constructing the NJ tree. Independent filtering of the original VCF based on the genotype missing rate, sequencing depth, and the number of allele types yielded 1,441,743 SNPs. After LD pruning, 265,365 SNPs were retained for PCA and admixture analyses.

Genetic clustering reveals distinct groups and broad correspondence with floral traits

After Sample 30-6 was pruned, the resulting NJ tree contained 167 *Hedychium* samples and resolved the major genetic relationships among the retained accessions (Fig. 1). The tree was rooted using *H. villosum* as a functional outgroup. This choice was supported by a previous molecular phylogenetic study of Thai *Hedychium*, in which *H. villosum* was separated from the group containing *H. coronarium*, *H. flavescens*, *H. coccineum*, and *H. forrestii*. Following the outgroup, *Hedychium* sp. (thick-leaved) was the most genetically distinct accession among the remaining samples, separating from all other ingroup accessions as an independent lineage. Unknown 3 diverged next and exhibited the longest branch length among all accessions, suggesting that it may represent a distinct genetic lineage not fully represented in this study, potentially a species not included in the current collection. The remaining 163 samples were further grouped into eight major clusters (Clades 1–8), a pattern highly consistent with the subsequent PCA and population structure analyses.

The clustering patterns showed a broad correspondence with several visible ornamental traits, especially flower color and floral architecture (Supplementary Fig. S2). Clades 1 and 2 predominantly consisted of germplasms with orange-red flowers. Specifically, Clade 1 was centered on *H. coccineum* and clustered with several USA cultivars ('Fireflies', 'Tara', and 'Orange Brush'). Their close genetic proximity, together with their shared orange-red coloration and small petal traits, suggests a close genetic affinity between *H. coccineum* and these cultivars. In contrast, Clade 8 comprised a group with white or pale-yellow flowers centered on *H. coronarium*, including closely related species (*H. flavescens* and *H. chrysoleucum*) and associated cultivars ('Baihuang Yu', 'Han Yue'), which generally feature large petals and imbricate bracts. These observations indicate a broad descriptive correspondence between genetic clustering and visible floral traits, particularly flower color and floral architecture.

Furthermore, the phylogenetic positions of several unknown accessions provided key clues for their identification: Unknown 1 clustered closely with the cultivar 'Tarissima', Unknown 2 grouped within the same branch as 'Elizabeth', and Unknown 4 showed a sister group relationship with *H. forrestii*. These results provide strong evidence for close genetic affinities.

A notable observation was the intermingled clustering of individuals from the cultivars 'Tara' and 'Orange Brush' within the phylogenetic tree, where they failed to form distinct monophyletic groups. This pattern, together with their highly similar morphological characteristics, indicates an extremely close genetic relationship between these two cultivars, potentially reflecting a closely related breeding origin or a recent shared genetic background.

Population genetic structure of *Hedychium* germplasm elucidated by PCA

PCA was performed using 265,365 LD-pruned SNPs after excluding *H. villosum* and Unknown 3, which showed extreme

Tree scale: 0.1

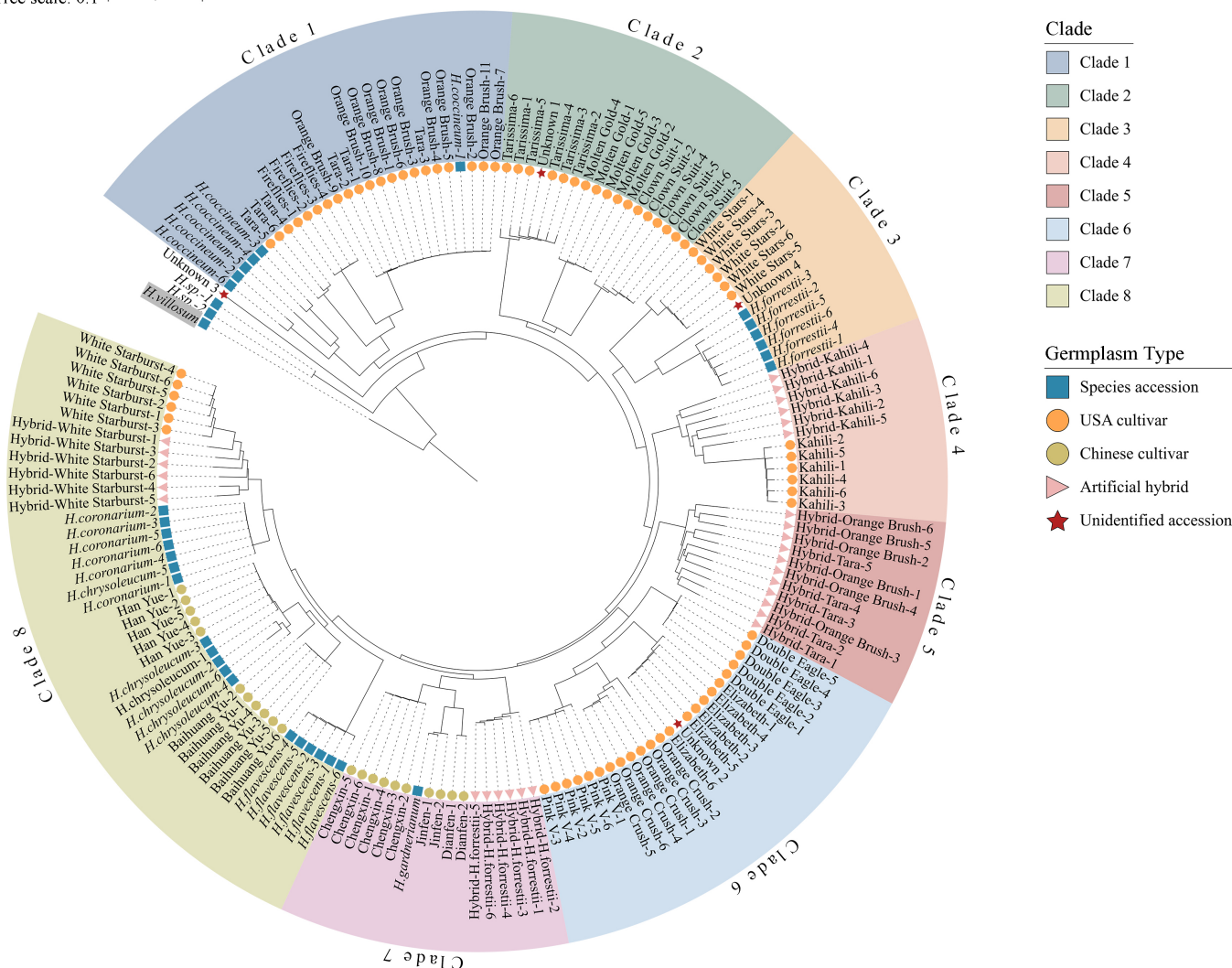


Fig. 1 Pruned NJ tree based on 647,693 high-quality SNPs. The original tree included 168 samples; Sample 30-6 was removed from the displayed topology, leaving 167 samples. The tree was rooted using *H. villosum* as a functional outgroup. Colors indicate the eight major clades.

divergence in the preliminary analysis of all 167 samples (Supplementary Fig. S1). In the final PCA of 165 samples, Principal Component (PC1) and PC2 explained 17.32% and 14.22% of the total genetic variation, respectively (Fig. 2). The PCA pattern largely supported the clustering pattern revealed by the NJ tree, with several major clades occupying distinguishable regions in the space defined by PC1 and PC2. In particular, Clades 1, 2, 3, 6, 7, and 8 showed clear separation to varying degrees, whereas Clades 4 and 5 were positioned close to each other near the center of the plot. *Hedychium* sp. (thick-leaved) occupied a clearly isolated position, further supporting its distinct genetic background relative to the other sampled accessions.

The PCA also provided informative evidence for hybrid and unknown accessions. The five artificial hybrid accessions were located between their corresponding parental groups, consistent with their expected hybrid origins. Unknown 1 almost overlapped with 'Tarissima', and Unknown 2 was positioned very close to 'Elizabeth', supporting their close genetic affinities with these cultivars. This genetic proximity between Unknown 2 and 'Elizabeth' is also consistent with their similar plant architecture and leaf morphology, as recorded in the living collection. Unknown 4 occupied a distinct

position, although *H. forrestii* was the closest sampled accession in the PCA space, suggesting a possible genetic affinity rather than complete identity. In addition, *H. gardnerianum* and the Chinese cultivar 'Jinfen' almost overlapped in the PCA plot, suggesting a very close genetic affinity between these two accessions.

Dissecting patterns of genetic admixture in *Hedychium* germplasm

Admixture analysis was performed using the 167 samples in the LD-pruned SNP dataset. The CV error generally decreased as K increased from 2 to 20 and reached its lowest value at K = 18 (Fig. 3). However, the profiles at high K values mainly reflected fine-scale subdivisions among closely related cultivars and replicate individuals rather than the major genetic groupings. Therefore, K = 8 was used as the main value for biological interpretation, because it showed strong concordance with the eight major clades identified by the NJ tree and with the PCA results.

At K = 8, most clades showed characteristic genetic component profiles, supporting the major clustering pattern revealed by the NJ tree (Fig. 4). Clade 4, consisting of 'Kahili' and hybrid 'Kahili', showed

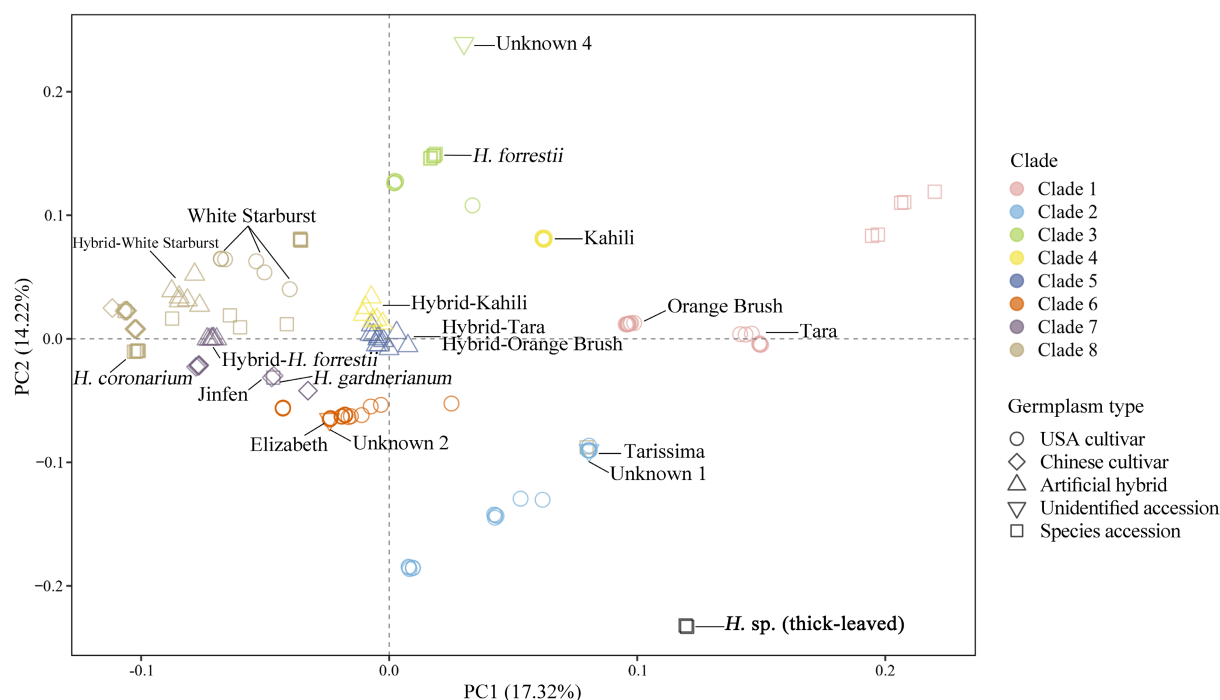


Fig. 2 PCA of 165 samples based on 265,365 SNPs after LD pruning. Colors indicate clades, and the shapes indicate germplasm types.

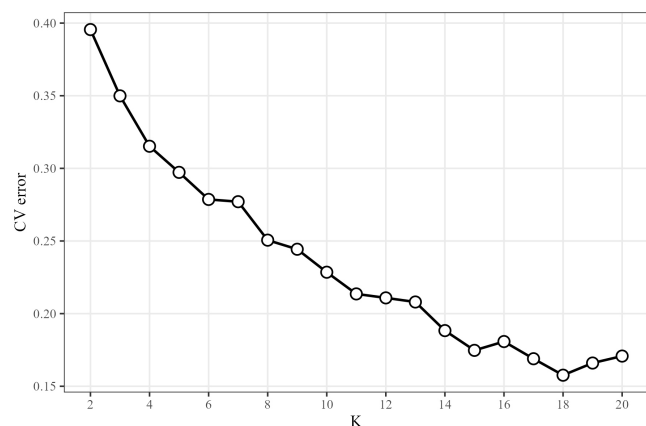


Fig. 3 CV error for different values of K.

a distinct pattern in which 'Kahili' was dominated by a single major component, whereas hybrid 'Kahili' displayed a mixed profile. Clade 5, composed of hybrid 'Tara' and hybrid 'Orange Brush', was also admixed, with two major components and minor contributions from other components. These patterns are consistent with the recorded hybrid backgrounds of these materials.

Clade 8 contained two major component patterns at $K = 8$. *H. coronarium* and materials related to 'White Starburst' were mainly characterized by one major component, whereas *H. flavescens* and 'Han Yue' were dominated by another component. 'Baihuang Yu' displayed a mixed profile between these two major components, whereas *H. chrysoleucum* showed a distinct combination of components. This pattern indicates that Clade 8 contains genetically differentiated but related white or pale-yellow flowered germplasm rather than a completely homogeneous group.

The admixture results also provided additional evidence for evaluating the genetic affinities of several unknown accessions. Unknown 1 and 'Tarissima' showed identical single-component profiles across all displayed K values, and Unknown 2 showed the same pattern as

'Elizabeth', supporting their very close genetic affinities with these cultivars. Unknown 4 exhibited the same component profile as *H. forrestii* at $K = 6$ to $K = 8$, whereas minor differences appeared at $K = 9$, $K = 10$, and $K = 18$. This pattern suggests a close genetic affinity between Unknown 4 and *H. forrestii*, while also indicating subtle differentiation at finer levels of population subdivision. In addition, *H. gardnerianum* and the Chinese cultivar 'Jinfen' exhibited highly similar component profiles, further supporting their close genetic affinity.

D-statistic analysis reveals asymmetric allele sharing among selected *Hedychium* accessions

D-statistic analysis was used as a complementary approach to examine asymmetric allele sharing among selected *Hedychium* accessions. Significant D-statistic signals ($|Z| > 3$) were detected in multiple combinations of species accessions and cultivars, indicating that the genetic relationships among these accessions cannot be fully explained by a strictly bifurcating tree model. Representative significant tests are summarized in Table 1, and the complete results are provided in Supplementary Tables S4 and S5.

In the representative cultivar tests, several cultivars showed significant asymmetric allele sharing with different species accessions or closely related lineages. For example, 'Molten Gold' showed significant signals in tests involving *H. gardnerianum* and *H. coronarium*, whereas 'Fireflies' showed significant allele sharing patterns involving both *H. coccineum* and comparisons involving *H. gardnerianum*. 'White Starburst' showed a significant signal in the comparison involving *H. coronarium*, consistent with its placement within the related *H. coronarium* group in the NJ tree and admixture analyses. These results indicate complex patterns of asymmetric allele sharing among several cultivars, species accessions, and closely related lineages. However, these D-statistic results were interpreted as evidence of asymmetric allele sharing rather than direct quantitative estimates of parental contribution.

D-statistic profile comparisons further supported the interpretation of several unknown accessions (Table 2). The detailed paired

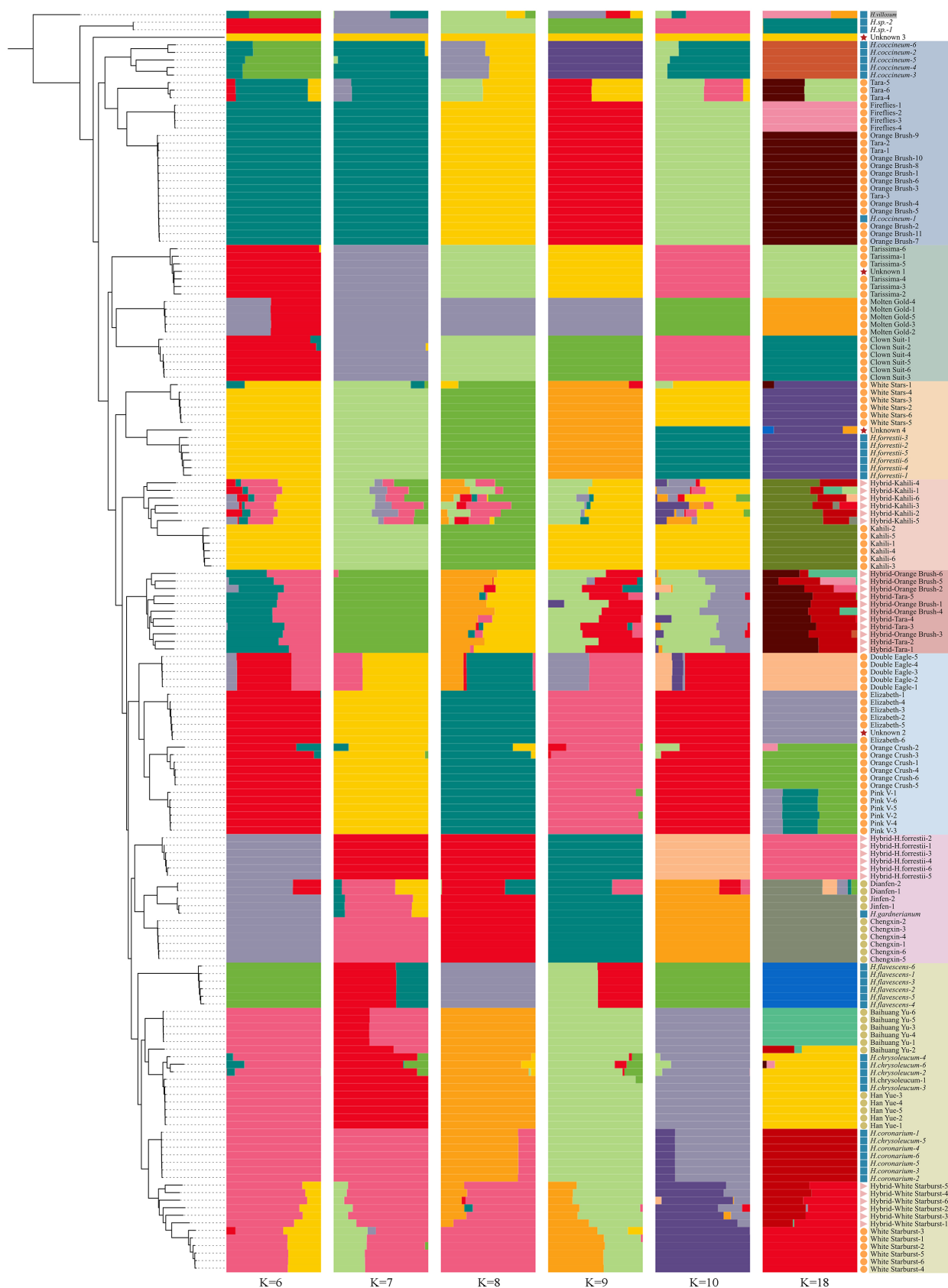


Fig. 4 Admixture results of the 167 *Hedychium* samples based on 265,365 SNPs after LD pruning. Results for K = 6, 7, 8, 9, 10, and 18 are shown. The sample order follows the NJ tree.

Table 1. Representative D-statistic tests showing significant asymmetric allele sharing among selected *Hedychium* accessions.

P1	P2	P3	D	Z-score	p-value
<i>H. coccineum</i>	<i>H. gardnerianum</i>	'Molten Gold'	0.719	20.652	< 0.001
<i>H. forrestii</i>	<i>H. coronarium</i>	'Molten Gold'	0.766	12.117	< 0.001
<i>H. coronarium</i>	<i>H. gardnerianum</i>	'Fireflies'	0.468	6.578	< 0.001
<i>H. coronarium</i>	<i>H. coccineum</i>	'Fireflies'	0.640	10.330	< 0.001
<i>H. chrysoleucum</i>	<i>H. coronarium</i>	'White Starburst'	0.356	6.933	< 0.001
<i>H. coronarium</i>	<i>H. gardnerianum</i>	'Orange Brush'	0.630	12.761	< 0.001
<i>H. coronarium</i>	<i>H. gardnerianum</i>	'Tara'	0.604	11.966	< 0.001
<i>H. forrestii</i>	<i>H. coronarium</i>	'Tarissima'	0.552	8.848	< 0.001

Note: Positive D-values indicate excess allele sharing between P3 and P2 relative to P3 and P1 under the reported Dsuite configuration. Significant results were defined as $|Z| > 3$. These results indicate significant asymmetric allele sharing and should not be interpreted as direct estimates of parental contribution.

Table 2. D-statistic profile comparisons between unknown accessions and their closest reference accessions.

Unknown accession	Compared reference	No. of paired tests	D profile correlation	Mean $ \Delta D $	Median $ \Delta D $
Unknown 1	'Tarissima'	441	0.983439	0.033351	0.027903
Unknown 2	'Elizabeth'	426	0.971047	0.027970	0.017218
Unknown 4	<i>H. forrestii</i>	415	0.966883	0.038125	0.028380

Note: Paired tests were defined as D-statistic tests in which the unknown accession and the compared reference accession occupied the same position in the four-taxon configuration, with the other two taxa identical and in the same order. D profile correlation represents the Pearson correlation between paired D-values. Mean $|\Delta D|$ and median $|\Delta D|$ indicate the absolute differences between paired D-values. These metrics were calculated from the complete D-statistic result set and were used as descriptive measures of D-statistic profile similarity rather than formal tests of accessions' identity. Detailed paired D-statistic comparisons for Unknown 1, Unknown 2, and Unknown 4 are provided in Supplementary Tables S6, S7, and S8, respectively.

comparisons underlying these D-statistic profile metrics are provided in [Supplementary Tables S6–S8](#). Unknown 1 showed a highly similar D-statistic profile to 'Tarissima', and Unknown 2 showed the strongest similarity to 'Elizabeth', as indicated by high profile correlations and low absolute differences in their paired D-values. Unknown 4 was also broadly similar to *H. forrestii*, but its lower profile correlation and larger absolute differences suggest greater differentiation than observed for Unknown 1 and Unknown 2. These results are consistent with the NJ tree, PCA, and admixture analyses, supporting close genetic affinities between Unknown 1 and 'Tarissima', between Unknown 2 and 'Elizabeth', and between Unknown 4 and *H. forrestii*, while indicating that Unknown 4 may represent a more differentiated accession.

Discussion

The genus *Hedychium* harbors rich genetic diversity and a long history of cultivation, yet the genetic relationships and breeding backgrounds of many modern cultivars remain poorly understood, partly because species delimitation in this genus has long been complicated by morphological similarity, interspecific variation, and hybridization^[4,5,9]. This study utilized SNP markers developed via RAD-seq to conduct a population genomic analysis of *Hedychium* germplasm representing 35 accessions, including major cultivated types, species accessions, artificial hybrids, and accessions of uncertain identity. Integrated analyses of their phylogenetic relationships, principal components, population structure, and D-statistics revealed a complex genetic architecture characterized by clear genetic differentiation, admixed component profiles, and asymmetric allele sharing among *Hedychium* germplasm. This work provides useful insights into the breeding history, germplasm classification, conservation, and future utilization of *Hedychium* genetic resources.

Admixture patterns associated with historical hybridization in modern *Hedychium* cultivars

Multiple lines of evidence from this study indicate that modern *Hedychium* cultivars possess complex genetic backgrounds in which

admixed genomic profiles are likely associated with historical hybridization during cultivation and breeding. The intermediate positions of artificial hybrids in the PCA and the mixed component profiles detected by admixture were consistent with their recorded hybrid backgrounds, whereas the significant asymmetric allele sharing signals detected by D-statistic analysis provided complementary evidence for genetic discordance among selected species accessions and cultivars. Together, these results suggest that historical hybridization and related genomic mixing may have contributed to the genetic structure of cultivated *Hedychium*. This interpretation is consistent with the known breeding history of many ornamental plants, in which breeders have widely used interspecific and inter-cultivar hybridization to combine desirable traits from different parents, such as flower color, form, fragrance, and adaptability^[29,30].

The admixed genetic profiles observed in several cultivars and artificial hybrids provide genomic evidence consistent with this breeding background. For example, 'Baihuang Yu' and several hybrid progenies showed mixed admixture component profiles, suggesting that they carry genetic components associated with more than one associated species or cultivated background. Widespread self-incompatibility and relatively high interspecific cross-compatibility in *Hedychium* may further facilitate hybridization under both natural and artificial conditions^[7].

Despite the admixture patterns associated with possible historical hybridization, the sampled germplasm also showed clear genetic differentiation among several major groups. For example, groups associated with *H. coccineum*, *H. forrestii* and *Hedychium* sp. (thick-leaved) retained relatively distinct genetic positions in the NJ tree, PCA, and admixture analyses. This pattern suggests that hybridization has not completely homogenized the genetic structure of cultivated *Hedychium*. Instead, modern germplasm appears to comprise both admixed cultivated materials and genetically differentiated backgrounds represented by the sampled species accessions.

Multiple species accessions associated with cultivated *Hedychium* germplasm

The integrated analyses in this study indicate that cultivated *Hedychium* germplasm shows genetic affinities with multiple

sampled species accessions or closely related taxa. The NJ tree, PCA, and admixture consistently revealed several distinct genetic groups, suggesting that modern cultivars were shaped by genetically diverse lineages rather than by a single ancestral source, and the D-statistic analysis provided complementary evidence of asymmetric allele sharing among selected accessions. This pattern is consistent with previous reports showing that hybridization has been used to combine desirable ornamental traits, such as floral scent, flower color, and floral morphology, in *Hedychium* breeding^[29].

Several representative species accessions showed clear associations with different cultivated groups. The group associated with *H. coronarium* was mainly represented in Clade 8, which included *H. coronarium*, closely related taxa such as *H. flavescens* and *H. chrysoleucum*, cultivars with white or pale yellow flowers, and several hybrid materials. The internal heterogeneity of this clade in the admixture analysis suggests that this cultivated group involved multiple closely related species accessions or taxa rather than a single homogeneous background. In contrast, *H. coccineum* was closely associated with Clade 1, which included cultivars with orange to red flowers, such as 'Tara', 'Orange Brush', and 'Fireflies', indicating its close relationship with this cultivar group. The group associated with *H. gardnerianum* was mainly represented in Clade 7 and several Chinese cultivated materials, especially 'Jinfen', which showed very close genetic affinity with *H. gardnerianum* in the PCA and admixture analyses. In addition, *H. forrestii* formed a genetically distinct group in Clade 3 together with 'White Stars' and Unknown 4, suggesting that it represents another differentiated lineage within the sampled germplasm.

D-statistic analysis provided complementary evidence for asymmetric allele sharing among several species accessions and cultivars, but these results were not treated as direct quantitative estimates of parental contribution, introgression proportion, or hybridization direction^[26–28]. Instead, they were interpreted together with the NJ tree, PCA, and admixture results. Taken together, these analyses suggest that modern *Hedychium* cultivars have complex and partially overlapping genetic backgrounds, with different species accessions or closely related taxa being associated with distinct genetic backgrounds in different cultivated groups.

Genetically differentiated accessions as potential resources for future breeding

In addition to the genetic backgrounds widely represented in cultivated groups, this study identified several species accessions and provisionally identified accessions that retain distinct genetic characteristics and may provide useful resources for future breeding, consistent with the broader importance of conserving and evaluating diverse *Hedychium* germplasm^[31,32]. Among them, *Hedychium* sp. (thick-leaved) showed the most pronounced differentiation among the sampled accessions, with its distinct position consistently supported by the NJ tree, PCA, and admixture results. No significant allele sharing signal involving this accession was detected in the tested D-statistic combinations.

H. forrestii also represented a relatively differentiated genetic background. It was assigned to Clade 3 together with Unknown 4 and 'White Stars', and this clade was separated from the major cultivar groups in the NJ tree and PCA plot. In the admixture analysis, Clade 3 showed a distinct component profile from $K = 8$ onward. The significant allele sharing signals involving *H. forrestii* were limited to a small number of cultivars, suggesting that its genetic contribution to the current cultivar set may be relatively restricted.

Other species accessions, such as *H. flavescens* and *H. chrysoleucum*, may contribute to the internal differentiation of Clade

8. Their admixture profiles were similar to those of some accessions with pale flowers, indicating that this clade contains multiple related but genetically differentiated backgrounds. These materials may therefore be useful for further diversification of white- or pale yellow-flowered germplasm.

These findings highlight the value of conserving and evaluating both genetically differentiated species accessions and provisionally identified *Hedychium* germplasm. Although systematic phenotypic evaluation was not conducted in this study, these materials may retain allelic variation that has not been fully captured in the current cultivated backgrounds. After further assessment of their ornamental traits, ecological adaptation, fertility, and cross-compatibility, they could provide useful variation for broadening the genetic base of cultivated *Hedychium*^[33]. Interspecific hybridization has been shown to improve floral scent, flower color, and morphological traits in *Hedychium*, and related strategies have also been applied in other ornamental or crop systems^[30,34–36]. Therefore, controlled interspecific hybridization, combined, where necessary, with embryo rescue, marker development, and marker-assisted selection, may provide a feasible route for introducing valuable variation(s) from genetically differentiated species accessions into future breeding programs^[37,38].

Genetic grouping, ornamental traits, and future perspectives

The genetic grouping identified in this study showed a broad correspondence with several visible ornamental traits, particularly flower color. For example, the clade associated with *H. coccineum* (Clade 1) consisted predominantly of accessions with orange to red flowers, whereas the group associated with *H. coronarium* (Clade 8) was mainly composed of accessions with white or pale yellow flowers. This pattern suggests that the use of genetically related parental backgrounds could have contributed to the clustering of cultivars with similar floral traits. However, because systematic phenotypic measurements and trait association tests were not conducted in this study, these observations should be interpreted as descriptive associations rather than direct evidence for the genetic control of ornamental traits.

The genome-wide SNP dataset generated in this study provides a useful foundation for future analyses of trait-related genetic variation in *Hedychium*. With expanded sampling and standardized phenotypic evaluation, this genomic framework could be used to investigate the genetic basis of key ornamental traits, such as flower color, floral scent, flowering time, and inflorescence architecture. Recent metabolomic, transcriptomic, and sensory omics studies in *Hedychium* have provided valuable insights into floral scent formation, volatile compound variation, and fragrance type differentiation, suggesting that population genomic resources could be integrated with multi-omics approaches in future trait studies^[39–41]. Future studies integrating population genomics with genome-wide association analysis, transcriptomics, metabolomics, and volatile compound profiling would be particularly valuable for identifying candidate loci and biological pathways related to ornamental diversification^[42–45]. Such integrative approaches may facilitate the transition of *Hedychium* breeding from phenotype-based selection toward more precise molecular breeding.

This study provides a population genomic framework for understanding the genetic differentiation, admixture patterns, and genetic affinities of *Hedychium* germplasm. The results indicate that modern cultivated materials are associated with multiple genetic backgrounds represented by the sampled species accessions or closely related taxa. Recorded artificial hybrids displayed mixed

genetic components consistent with their documented origins, whereas several cultivars showed admixture patterns potentially associated with historical hybridization. Several genetically differentiated accessions, including the provisionally identified *Hedychium* sp. (thick-leaved) and *H. forrestii*, may represent useful resources for broadening the genetic base of future breeding after further phenotypic and reproductive evaluation. These findings provide useful information for germplasm classification, conservation, and breeding applications in *Hedychium*.

Certain limitations of this study should be noted. First, RAD-seq is a reduced representation sequencing method based on restriction enzyme digestion, which may not capture variants in genomic regions lacking the targeted restriction sites and may be affected by missing data or sampling bias in some genomic regions^[15,46]. Future studies could utilize whole-genome resequencing to obtain more comprehensive genomic information^[47–49]. Second, the NJ tree was rooted using *H. villosum* as a functional outgroup within the genus rather than a true external outgroup from a related Zingiberaceae genus; incorporating such data in future studies would further strengthen the phylogenetic framework. Third, the present study focused primarily on genomic relationships and did not include systematic phenotypic, cytological, or reproductive compatibility assessments. Integrating these data in future research would allow a more direct evaluation of the relationships among genetic background, ornamental traits, ploidy variation, and breeding potential^[50,51]. Additionally, our samples were primarily from China and the USA, with limited representation of *Hedychium* species diversity from the Himalayan region and other parts of tropical and subtropical Asia. Expanding the collection to include a wider geographical range and more diverse sample types would offer a more complete depiction of the genetic landscape of *Hedychium*.

Conclusions

This study characterized the population genetic structure of *Hedychium* germplasm using RAD-seq. After Sample 30-6 was removed, the displayed NJ topology contained 167 samples. The admixture and D-statistic analyses were rerun using 167-sample datasets, and the final PCA included 165 samples after the exclusion of two extreme outliers. The NJ tree resolved eight major clades, which were broadly supported by the PCA and admixture results. D-statistic analysis detected asymmetric allele sharing in selected comparisons involving species accessions and cultivars, indicating complex and partially overlapping genetic backgrounds.

Modern cultivated materials showed genetic affinities with multiple backgrounds represented by the sampled species accessions, particularly those related to *H. coronarium*, *H. coccineum*, *H. gardnerianum*, and *H. forrestii*. Recorded artificial hybrids generally exhibited mixed genetic components consistent with their recorded origins, whereas several cultivars showed admixture patterns potentially associated with historical hybridization. Additionally, several genetically differentiated accessions, such as *Hedychium* sp. (thick-leaved) and *H. forrestii*, may provide useful genetic variation for broadening the genetic base of future breeding after further phenotypic and reproductive evaluation. Overall, this research establishes a population genomic framework for germplasm classification, conservation, and molecular breeding in *Hedychium*.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Cai B, Bao P; writing – original draft,

reviewing, and editing: Luo X, Cai B; investigation, data analysis, and supervision: Luo X, Bao P, Cai C, Wang M, Xie Y, Wang J; funding acquisition: Cai B. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The raw RAD-seq sequencing reads have been deposited in the NCBI SRA under BioProject accession PRJNA1484304, and the Data Availability section has been updated accordingly. As the data have not yet been publicly released, the deposited submission information can be viewed using the following NCBI reviewer link: <https://dataview.ncbi.nlm.nih.gov/object/PRJNA1484304?reviewer=1jpnh7p4kkb2dg06puul4gsbu>.

Acknowledgments

This research was supported by the Social Development Guiding Projects of Xiamen City, China (No. 3502ZZ20234ZD4002).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information This accompanies this paper online at: <https://doi.org/10.48130/opr-0026-0026>.

Dates

Received 14 February 2026; Revised 25 June 2026; Accepted 3 July 2026; Published online 9 October 2026

References

- [1] Wongsuwan P, Pichansoonthon C. 2011. Taxonomic revision of the genus *Hedychium* J. König (Zingiberaceae) in Thailand (Part I). *The Journal of the Royal Institute of Thailand* III: 126–149 (in Thai)
- [2] Wu D, Larsen K. 2000. Zingiberaceae. In *Flora of China*, eds. ZY Wu, PH Raven. Vol. 24. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. pp. 322–377 www.iplant.cn/foc/pdf/Zingiberaceae.pdf
- [3] Zhou Y, Wang Q, Wei X, Gao T, Abbas F, et al. 2024. Dynamic aromatics: Evaluating fragrance quality shifts and implementation of real-time rapid detection in *Hedychium* cut flowers during senescence. *Postharvest Biology and Technology* 218:113193
- [4] Wood TH, Whitten WM, Williams NH. 2000. Phylogeny of *Hedychium* and related genera (Zingiberaceae) based on ITS sequence data. *Edinburgh Journal of Botany* 57:261–270
- [5] Saryan P, Gupta S, Gowda V. 2020. Species complex delimitations in the genus *Hedychium*: a machine learning approach for cluster discovery. *Applications in Plant Sciences* 8:e11377
- [6] Warlarphih D, Suchiang W, Susngi AM, Lamo JM. 2022. Genetic diversity and species relationship of *Hedychium* J. Koenig as revealed by DAMD and ISJ markers. *The Nucleus* 65:359–367
- [7] Gao L, Liu N, Huang B, Hu X. 2008. Phylogenetic analysis and genetic mapping of Chinese *Hedychium* using SRAP markers. *Scientia Horticulturae* 117:369–377
- [8] Jugran A, Bhatt ID, Rawat S, Giri L, Rawal RS, et al. 2011. Genetic diversity and differentiation in *Hedychium spicatum*, a valuable medicinal plant of Indian Himalaya. *Biochemical Genetics* 49:806–818
- [9] Suksathan R, Anuntalabhochai S, Jampeetong A, Sookkhee S, Chansakaow S. 2014. A phylogenetic analysis of Thai *Hedychium* (Zingiberaceae) and development of SCAR marker for *Hedychium flavescens* Carey ex Roscoe. *Chiang Mai Journal of Science* 41:286–297
- [10] Basak S, Ramesh AM, Kesari V, Parida A, Mitra S, et al. 2014. Genetic diversity and relationship of *Hedychium* from Northeast India as dissected using PCA analysis and hierarchical clustering. *Meta Gene* 2:459–468

- [11] Rajkumari S, Sanatombi K. 2019. Genetic diversity analysis of *Hedychium* species based on RAPD and ISSR markers. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences* 89:623–629
- [12] Zhou Y, Wei X, Abbas F, Yu Y, Yu R, et al. 2021. Genome-wide identification of simple sequence repeats and assessment of genetic diversity in *Hedychium*. *Journal of Applied Research on Medicinal and Aromatic Plants* 24:100312
- [13] Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* 17:240–248
- [14] Guo C, Ma PF, Yang GQ, Ye XY, Guo Y, et al. 2021. Parallel ddRAD and genome skimming analyses reveal a radiative and reticulate evolutionary history of the temperate bamboos. *Systematic Biology* 70:756–773
- [15] Rivera-Colón AG, Rochette NC, Catchen JM. 2021. Simulation with RADinitio improves RADseq experimental design and sheds light on sources of missing data. *Molecular Ecology Resources* 21:363–378
- [16] Heckenhauer J, Samuel R, Ashton PS, Abu Salim K, Paun O. 2018. Phylogenomics resolves evolutionary relationships and provides insights into floral evolution in the tribe Shoreeae (Dipterocarpaceae). *Molecular Phylogenetics and Evolution* 127:1–13
- [17] Chang YH, Yao G, Neilsen J, Liu DT, Zhang L, et al. 2021. *Rhododendron kuomeianum* (Ericaceae), a new species from northeastern Yunnan (China), based on morphological and genomic data. *Plant Diversity* 43:292–298
- [18] Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3:e3376
- [19] Rochette NC, Rivera-Colón AG, Catchen JM. 2019. Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology* 28:4737–4754
- [20] Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22:3124–3140
- [21] Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, et al. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7
- [22] Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* 52:W78–W82
- [23] Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. 2nd Edition. Cham: Springer. doi: [10.1007/978-3-319-24277-4](https://doi.org/10.1007/978-3-319-24277-4)
- [24] Alexander DH, Lange K. 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12:246
- [25] Chen C, Wu Y, Li J, Wang X, Zeng Z, et al. 2023. TBtools-II: a 'one for all, all for one' bioinformatics platform for biological big-data mining. *Molecular Plant* 16:1733–1742
- [26] Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources* 21:584–595
- [27] Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution* 28:2239–2252
- [28] Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Molecular Biology and Evolution* 32:244–257
- [29] Wei X, Zhou Y, Abbas F, Yan F, Zou X, et al. 2023. Distant heteroploid hybridization improved *Hedychium* floral scent, floral color and morphological traits. *Industrial Crops and Products* 194:116357
- [30] Kuligowska K, Lütken H, Müller R. 2016. Towards development of new ornamental plants: status and progress in wide hybridization. *Planta* 244:1–17
- [31] Sarangthem N, Talukdar NC, Thongam B. 2013. Collection and evaluation of *Hedychium* species of Manipur, northeast India. *Genetic Resources and Crop Evolution* 60:13–21
- [32] Cheng B, Zhao K, Zhou M, Bourke PM, Zhou L, et al. 2025. Phenotypic and genomic signatures across wild *Rosa* species open new horizons for modern rose breeding. *Nature Plants* 11:775–789
- [33] Gao JY, Liu Q, Li QJ. 2014. The comparative reproductive biology of a tetraploid species, *Hedychium villosum*, and its diploid progenitor *H. tenuiflorum* (Zingiberaceae). *Plant Biology* 16:683–689
- [34] Song R, Zhang D, Yang J, Cheng Y, Song X, et al. 2024. Identification and transferring of a new *Fusarium* head blight resistance gene *FhbRc2* from *Roegneria ciliaris* 3S^L chromosome arm into common wheat. *The Crop Journal* 12:1718–1726
- [35] Lin Y, Zhou S, Liang X, Han B, Yang J, et al. 2023. Introgression of chromosome 6PL terminal segment from *Agropyron cristatum* to increase both grain number and grain weight in wheat. *The Crop Journal* 11:878–886
- [36] Lin Q, Wang X, Wan Z, Tao J, Chi X, et al. 2025. Obtained intergeneric hybrids from distant hybridization between *Lagerstroemia* and *Heimia*. *Ornamental Plant Research* 5:e017
- [37] Yang Y, Sun M, Li S, Chen Q, Teixeira da Silva JA, et al. 2020. Germplasm resources and genetic breeding of *Paonia*: a systematic review. *Horticulture Research* 7:107
- [38] He J, Liu S, Zhong S, Chen Y, Luo Y, et al. 2026. Germplasm resource and genetic breeding of ornamental Madagascar periwinkle (*Catharanthus roseus*): a brief review. *Ornamental Plant Research* 6:e003
- [39] Zhou Y, Abbas F, Wang Z, Yu Y, Yue Y, et al. 2021. HS-SPME-GC-MS and electronic nose reveal differences in the volatile profiles of *Hedychium* flowers. *Molecules* 26:5425
- [40] Abbas F, Zhou Y, He J, Ke Y, Qin W, et al. 2021. Metabolite and transcriptome profiling analysis revealed that melatonin positively regulates floral scent production in *Hedychium coronarium*. *Frontiers in Plant Science* 12:808899
- [41] Zhou Y, Wang F, Wei X, Xiong M, Gao T, et al. 2026. From fragrance wheel to functional genes: a multi-omics investigation into fragrance type formation in ornamental *Hedychium* flowers. *Horticulture Research* 13(6):uhag063
- [42] Chen H, Zeng X, Yang J, Cai X, Shi Y, et al. 2021. Whole-genome resequencing of *Osmanthus fragrans* provides insights into flower color evolution. *Horticulture Research* 8:98
- [43] Zhou Y, Zheng T, Cai M, Feng L, Chi X, et al. 2023. Genome assembly and resequencing analyses provide new insights into the evolution, domestication and ornamental traits of crape myrtle. *Horticulture Research* 10:uhad146
- [44] Ye Q, Zhang L, Li Q, Ji Y, Zhou Y, et al. 2024. Genome and GWAS analysis identified genes significantly related to phenotypic state of *Rhododendron* bark. *Horticulture Research* 11:uhae008
- [45] Cheng B, Du W, Bourke PM, Yu C. 2024. Population genetics of horticultural crops aided by multi-omics technology and its implications for ornamental plants. *Ornamental Plant Research* 4:e024
- [46] Roure B, Baurain D, Philippe H. 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Molecular Biology and Evolution* 30:197–214
- [47] Dong Q, Zou QC, Mao LH, Tian DQ, Hu W, et al. 2022. The chromosome-scale assembly of the *Curcuma alismatifolia* genome provides insight into anthocyanin and terpenoid biosynthesis. *Frontiers in Plant Science* 13:899588
- [48] Hibrand Saint-Oyant L, Ruttink T, Hamama L, Kirov I, Lakhwani D, et al. 2018. A high-quality genome sequence of *Rosa chinensis* to elucidate ornamental traits. *Nature Plants* 4:473–484
- [49] Zhang Q, Zhang H, Sun L, Fan G, Ye M, et al. 2018. The genetic architecture of floral traits in the woody plant *Prunus mume*. *Nature Communications* 9:1702
- [50] Zhang X, Lan L, Yang Y, Guan H, Peng D, et al. 2026. Pangenomic analyses of rose uncover widespread structure variation and empower genomics-directed breeding. *Nature Genetics* 58:1164–1175
- [51] Harmon DD, Chen H, Byrne D, Liu W, Ranney TG. 2023. Cytogenetics, ploidy, and genome sizes of rose (*Rosa* spp.). *cultivars and breeding lines*. *Ornamental Plant Research* 3:10



Copyright: © 2026 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.