

Research paper

Cryptic diversity and rampant hybridization in annual gentians on the Qinghai-Tibet Plateau revealed by population genomic analysis


 Peng-Cheng Fu ^{a, *}, Qiao-Qiao Guo ^a, Di Chang ^a, Qing-Bo Gao ^b, Shan-Shan Sun ^a
^a School of Life Science, Luoyang Normal University, Luoyang 471934, PR China

^b Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810001, PR China

ARTICLE INFO

Article history:

Received 26 April 2023

Received in revised form

24 October 2023

Accepted 25 October 2023

Available online 28 October 2023

Keywords:

Gentiana aristata

Hybrid zone

Qinghai-Tibet Plateau

Plastome

Nuclear SNPs

ABSTRACT

Understanding the evolutionary and ecological processes involved in population differentiation and speciation provides critical insights into biodiversity formation. In this study, we employed 29,865 single nucleotide polymorphisms (SNPs) and complete plastomes to examine genomic divergence and hybridization in *Gentiana aristata*, which is endemic to the Qinghai-Tibet Plateau (QTP) region. Genetic clustering revealed that *G. aristata* is characterized by geographic genetic structures with five clusters (West, East, Central, South and North). The West cluster has a specific morphological character (i.e., blue corolla) and higher values of F_{ST} compared to the remaining clusters, likely the result of the geological barrier formed by the Yangtze River. The West cluster diverged from the other clusters in the Early Pliocene; these remaining clusters diverged from one another in the Early Quaternary. Phylogenetic reconstructions based on SNPs and plastid data revealed substantial cyto-nuclear conflicts. Genetic clustering and *D*-statistics demonstrated rampant hybridization between the Central and North clusters, along the Bayankala Mountains, which form the geological barrier between the Central and North clusters. Species distribution modeling demonstrated the range of *G. aristata* expanded since the Last Interglacial period. Our findings provide genetic and morphological evidence of cryptic diversity in *G. aristata*, and identified rampant hybridization between genetic clusters along a geological barrier. These findings suggest that geological barriers and climatic fluctuations have an important role in triggering diversification as well as hybridization, indicating that cryptic diversity and hybridization are essential factors in biodiversity formation within the QTP region.

Copyright © 2023 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The major mountain systems of the world are remarkably rich in species diversity (Antonelli et al., 2018), with alpine species being especially noteworthy. Alpine floras have been profoundly impacted by recurrent cycles of historical climatic change (Ding et al., 2020) and continue to be affected by current climate warming trends (Elsen and Tingley, 2015; Liang et al., 2018). These climatic oscillations have potentially caused isolation and reconnection between adjacent populations, with the resultant range fragmentation, range shifts, secondary contacts, and

hybridization accelerating speciation events (Muellner-Riehl et al., 2019; Kirschner et al., 2020). Frequent geographical isolation and increased levels of environmental heterogeneity may accelerate the diversity of cryptic species, which might either be a consequence of *in situ* speciation or colonization by lowland taxa followed by allopatric divergence (Struck et al., 2018). Although relevant to all mountain systems of the world, research has been principally focused on a few mountainous model regions, such as the Qinghai-Tibet Plateau (QTP) region.

The QTP region, including the Tibetan Plateau, Hengduan Mountains (HM), and the Himalayas (according to the geological definition by Liu et al. (2022)), comprises two global hotspots of biodiversity (Myers et al., 2000; Marchese, 2015). It is an ideal area for evaluating the spatial–temporal evolution of alpine communities as well as the factors that contribute to diversification and speciation. This region harbors a rich flora with a high proportion of

* Corresponding author.

E-mail address: fupengc@sina.com (P.-C. Fu).

Peer review under responsibility of Editorial Office of Plant Diversity.

endemic species and is characterized by a high rate of *in situ* speciation (Xing and Ree, 2017; Ding et al., 2020). For instance, the HM hosts particularly rich alpine flora comprising a diverse mosaic of floristic elements (Li et al., 2021). The QTP region, particularly the HM, is characterized by distinct geographic barriers and frequent climatic fluctuations (Favre et al., 2015; Sun et al., 2017), which subdivides distribution ranges and generates fine-scale environmental heterogeneity, therefore, stimulating divergence and speciation (Sun et al., 2017; Muellner-Riehl, 2019).

Besides geological barriers and climatic changes, biological processes, such as hybridization, have been essential in driving species formation and biodiversity assembly (Abbott et al., 2013; van Wyk et al., 2017). Hybrid speciation has been detected in substantial plant lineages around the world (Mallet, 2007; Rieseberg and Willis, 2007; Schumer et al., 2014), and dynamic hybrid zones are frequently observed in various lineages worldwide (Abbott et al., 2017). The major impact of global climate change on the dynamics and distribution of biodiversity presents an unprecedented opportunity to study hybridization, enabling us to peer into the birth and death of evolutionary lineages (Vallejo-Marín and Hiscock, 2016). In the QTP region, natural hybrids have been documented in almost all species-rich plant genera, and gene flow has been widely detected during the speciation processes of all taxonomic groups examined (Wu et al., 2022). This region has been suggested to be a site of dynamic flora as numerous species are presently becoming reproductively isolated (Wu et al., 2022). For example, hybridization and young speciation are frequently observed among not only speciose alpine herbs or shrubs, including *Gentiana* (Fu et al., 2020, 2022a), *Saxifraga* (Ebersbach et al., 2020) and *Rhododendron* (Zheng et al., 2021; Ma et al., 2022), but also woody species such as *Pterocarya* (Wang et al., 2023), *Populus* (Li et al., 2023) and *Betula* (Nocchi et al., 2023).

Gentiana (Gentianaceae) is an alpine genus with a global range. The QTP region serves as both the center of biodiversity for approximately 360 *Gentiana* species (Ho and Liu, 2001) and the primary source for colonization of other regions (Favre et al., 2016). As observed in the majority of plant lineages, abiotic factors, including geological processes, geographical isolation, and climatic changes, are correlated with diversification and speciation in several lineages of *Gentiana* (Zhang et al., 2009; Favre et al., 2016; Fu et al., 2018, 2020, 2022b; Sun et al., 2022). In addition, according to growing evidence from phylogenetic and population genetic studies, hybridization is frequently observed in *Gentiana*, both in the QTP region (Chen et al., 2021; Fu et al., 2021a, 2022a) and Europe (Favre et al., 2021). In *Gentiana*, section *Chondrophyllae* Bunge sensu lato (s.l.) is the only section that is globally distributed, which has 182 species and accounts for 51.7% of all *Gentiana* species (Ho and Liu, 2001; Favre et al., 2020). This section has experienced rapid evolution (Yuan and Küpfer, 1997) and presented extremely long branches in the phylogenetic trees (Favre et al., 2020; Chen et al., 2021; Fu et al., 2021b, 2022c). However, the reasons underlying the rich diversity in this section are yet unclear. Additionally, section *Chondrophyllae* s.l. is a well-supported monophyletic group, but the intra-sectional relationship is ambiguous as a result of rich cyto-nuclear conflicts (Chen et al., 2021; Fu et al., 2022c). Therefore, hybridization is assumed to be rampant in this section and has been essential in the assembly of biodiversity, although there is little direct evidence.

In this study, we focused on an annual gentian species *Gentiana aristata* Maxim. in *G.* section *Chondrophylla* s.l. *G. aristata* is endemic to the eastern QTP and occurs over a large range from the Qilian Mountains to the Hengduan Mountains (Ho and Pringle, 1995). It varies in morphological characteristics, for example, having purplish red or blue in the corolla (Ho and Liu, 2001). It possesses 12–15 mm corolla, with pollinators including bumblebees, flies and

ants (personal observation). We used genomic data to examine the fine-scale phylogeographic structures and evolutionary history of the species. Specifically, we asked (i) whether geographical features in the QTP region have acted as barriers to gene flow, leading to discrete genetic structures and the potential presence of cryptic lineages; and (ii) whether modifications of distribution ranges during climate oscillations have resulted in hybridization and possible hybrid zone.

2. Materials and methods

2.1. Sampling, library construction, and sequencing

A total of 133 individuals of *Gentiana aristata* from 23 different locations across its range were sampled (Table 1 and Fig. 1). Because the plant is tiny, the entire plant of each individual was collected and dried in silica gel. Our field work determined that the corolla of *G. aristata* is blue (Type B) in two locations (YS and LWQ) and purplish (Type A) in all others (Fig. 1). Voucher specimens were deposited in the herbarium of Luoyang Normal University.

Total genomic DNA was extracted from dried leaves with a Dzap plant genomic DNA extraction kit (Sangon, Shanghai, China). DNA concentrations were quantified using Qubit 2.0 fluorometer (Life Technologies). For Genotyping-By-Sequencing (GBS) library construction, a GBS pre-design experiment was performed to evaluate the enzymes and sizes of restriction fragments. We imposed three criteria: (i) the number of tags must be larger than 0.3 million; (ii) the enzymatic tags must be evenly distributed across the sequences to be examined; and (iii) repeated tags must be avoided. To maintain the homogeneity of sequence depth across different fragments, a small length range was selected (about 50 bp). Each sample was digested with restriction enzyme *MseI* (New England Biolabs, NEB), and thereafter with *NlaIII* (NEB) and *EcoRI* (NEB). The ligation products were subsequently purified and amplified with PCR. Fragments between 350 and 400 bp were isolated using a Gel Extraction Kit (Qiagen) and purified using Agencourt AMPure XP (Beckman). Libraries were multiplexed and sequenced using a 2 × 150 bp run on two lanes of Illumina NovaSeq 6000 (Tianjin, China).

To obtain plastome sequences, 23 samples representing all locations were selected to perform genome-skimming sequencing. For each sample, a 500-bp DNA Illumina NovaSeq sequencing library was constructed using approximately 4.0 ng of sonicated DNA as input. The libraries were multiplexed and sequenced using a 2 × 150 bp run on one lane of Illumina NovaSeq 6000 (Tianjin, China).

2.2. Processing of illumina data

Raw reads were filtered and trimmed using Trimmomatic 0.32 (Bolger et al., 2014) with default parameters to eliminate adaptor sequences and low-quality reads and sites. For GBS sequencing, the clean reads against the chromosome-level genome of *Gentiana dahurica* were mapped (PRJNA799480; Li et al., 2022); the closest available genome was mapped with bwa-mem v.2.2.1 (Li and Durbin, 2009), and the sequence alignment/map format files were produced employing SAMtools v.1.9 (Li et al., 2009). PCR duplications were identified using sambamba v.0.8.1 (Tarasov et al., 2015), and variations were assessed with freebayes v.0.9.21 (Garrison and Marth, 2012). Only SNPs were maintained in vcftools v.0.1.13 (Danecek et al., 2011) for downstream analysis. SNPs having a minor allele frequency (MAF) of less than 5% and missing frequency of more than 0.8 among individuals were filtered using vcftools v.0.1.13 (Danecek et al., 2011). Linkage-disequilibrium (LD) SNP pruning was performed in vcftools to eliminate variants from

Table 1

Genetic statistics for *Gentiana aristata*. No., sample size; Ar, allelic richness; H_s , mean observed gene diversities within population; H_o , mean observed heterozygosities within population; F_{IS} , mean inbreeding coefficient. Abbreviation after localities indicates provinces as follows: QH, Qinghai; GS, Gansu; SC, Sichuan; XZ, Xizang.

Group	ID	No.	Voucher Ref.	Location	Latitude (N)	Longitude (E)	Elevation (m/a.s.l)	Ar	H_o	H_s	F_{IS}	
North	XH	6	Fu2016001	Xunhua, QH	35°35'	102°42'	3094	1.2192	0.1437	0.2315	0.3793	
	LQ1	4	Fu2016016	Lvqu, GS	34°42'	102°29'	3281	1.2162	0.1522	0.2326	0.3453	
	GD	6	Fu2017352	Guide, QH	36°19'	101°29'	3600	1.2064	0.1294	0.2190	0.4090	
	LQ2	6	Fu2018016	Lvqu, GS	34°27'	102°18'	3603	1.2226	0.1373	0.2349	0.4156	
	MY1	6	Gao2019139	Mengyuan, QH	37°07'	102°27'	2978	1.2024	0.1272	0.2151	0.4086	
	MY2	6	Gao2019557	Mengyuan, QH	37°47'	101°19'	3640	1.2042	0.1359	0.2151	0.3683	
	QL1	6	Gao2019677	Qilian, QH	38°00'	100°21'	3209	1.1995	0.1257	0.2122	0.4075	
	QL2	6	Gao2019688	Qilian, QH	37°35'	100°45'	3439	1.2062	0.1284	0.2187	0.4129	
	Hybrids	GH	6	Fu2017008	Gonghe, QH	35°52'	99°58'	3360	1.2244	0.1518	0.2346	0.3528
		JZ1	6	Fu2017206	Jiuzhi, QH	33°12'	101°28'	3717	1.2197	0.1331	0.2325	0.4275
JZ2		6	Fu2017240	Jiuzhi, QH	33°27'	100°59'	3926	1.2353	0.1467	0.2478	0.4079	
MQ		6	Fu2017276	Maqin, QH	34°18'	100°16'	3985	1.2290	0.1455	0.2410	0.3962	
DR		6	Fu2017250	Dari, QH	33°39'	99°46'	4152	1.2223	0.1360	0.2353	0.4219	
Central	CD	6	Fu2017031	Chenduo, QH	33°07'	97°27'	4119	1.2229	0.1578	0.2330	0.3227	
	ZD	6	Fu2023049	Zaduo, QH	32°55'	95°38'	4559	1.1827	0.1040	0.2037	0.4895	
	QML	6	Fu2023059	Qumalai, QH	33°58'	96°35'	4594	1.1901	0.1042	0.2118	0.5080	
	BM	6	Fu2023073	Banma, QH	32°42'	100°40'	4420	1.1989	0.1046	0.2217	0.5283	
East	MEK	6	Fu2017182	Maerkang, SC	31°53'	102°39'	3343	1.1768	0.1006	0.1969	0.4892	
	HY	5	Fu2017195	Hongyuan, SC	32°09'	102°30'	3516	1.2084	0.1385	0.2235	0.3802	
South	LH	6	Fu2017162	Luhuo, SC	31°44'	100°43'	4022	1.1972	0.1180	0.2132	0.4465	
	KD	4	Fu2023011	Kangding, SC	30°04'	101°25'	3523	1.1054	0.0662	0.1342	0.5067	
West	YS	6	Fu2017055	Yushu, QH	32°46'	97°12'	4013	1.1394	0.0974	0.1548	0.3708	
	LWQ	6	Fu2017112	Leiwuqi, XZ	31°04'	96°56'	4626	1.1473	0.0997	0.1637	0.3908	

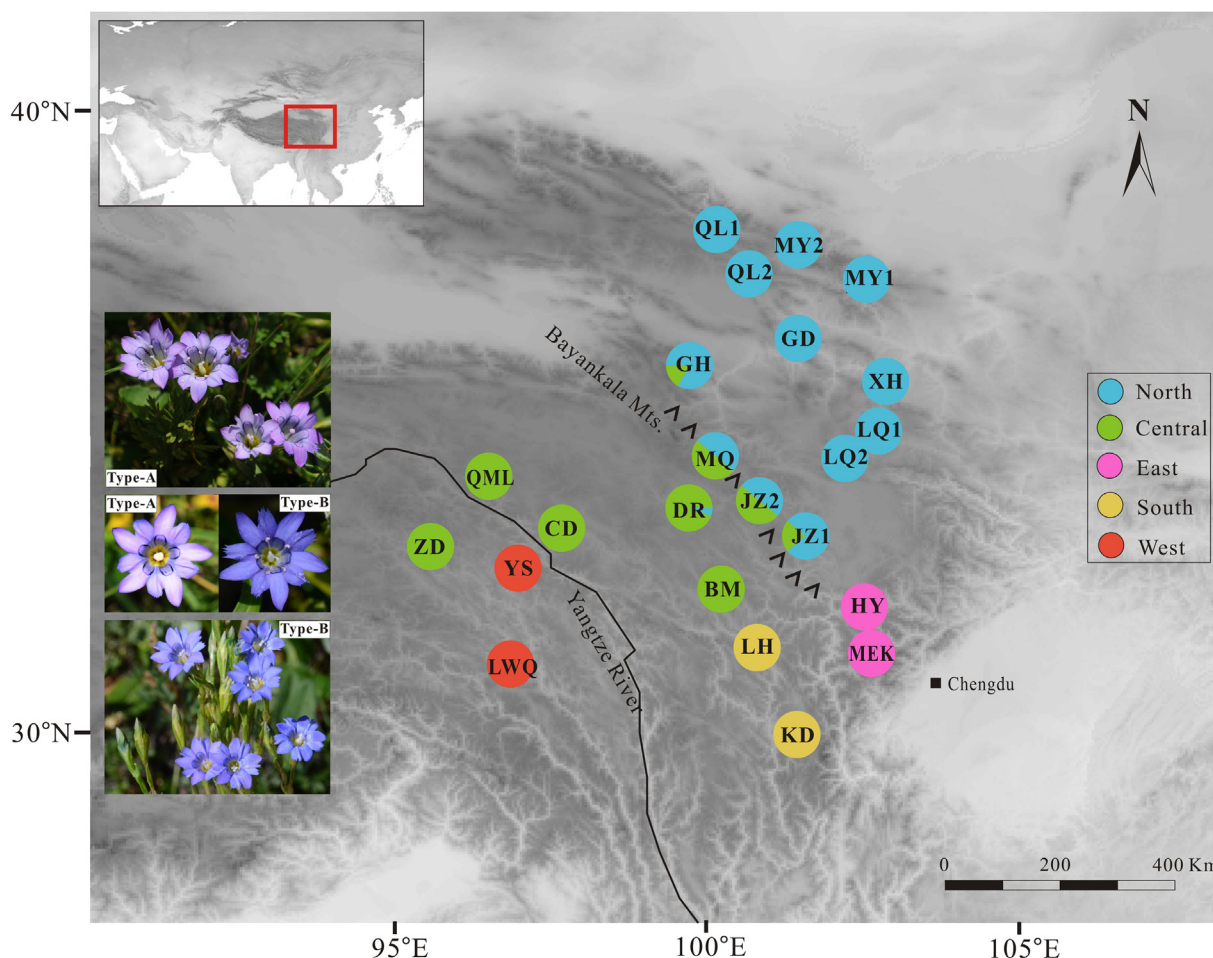


Fig. 1. Genetic structure of *Gentiana aristata* in the Qinghai-Tibet Plateau region. Pie charts illustrate the frequency of color-coded groups based on 29,865 SNPs in FastStructure. The five colors represent the five genetic clusters in Fig. 2. The photos reveal the color variation in the corolla. Type A has a purplish corolla and occurs in all *G. aristata* populations except YS and LWQ. Type B has a blue corolla and occurs in populations YS and LWQ. Photographs: Peng-Cheng Fu.

each pair that were closer than 100 bp. The final VCF file was converted into different formats using PGDSpider v.2.1.1.5 (Lischer and Excoffier, 2012) to perform further analyses.

Clean reads were assembled for genome-skimming sequencing using the GetOrganelle pipeline (Jin et al., 2020) with default parameters. The published plastome of *Gentiana aristata* (MN234139, Fu et al., 2021b) was used as the reference. Each plastid genome was annotated with Plastid Genome Annotator (PGA) (Qu et al., 2019) using the default parameters.

2.3. Genetic differentiation and population genetic structure

Allelic richness (A_r) was computed, and heterozygosity (H_0), gene diversity (H_S), and Wright's inbreeding coefficient (F_{IS}) were measured using the R package HierFstat (Goudet, 2005). To examine the degree of genetic differentiation between populations, pairwise F_{ST} was estimated by the Weir and Cockerham method (Weir and Cockerham, 1984) using the HierFstat package (Goudet, 2005). Range-wide isolation by distance (IBD) (Wright, 1943) was determined with a Mantel test using the geographic distance and pairwise genetic distance. $F_{ST}/(1-F_{ST})$ against pairwise geographic distances was plotted employing R v.4.0.1 (R Core Team, 2020).

To identify genetic clusters of *Gentiana aristata*, a Bayesian clustering method implemented in FastSTRUCTURE (Raj et al., 2014) was employed based on nuclear SNPs, with assumed clusters (K) from 1 to 18. The *chooseK.py* script was used to assess the model complexity for the data. Graphical representation of individual cluster assignments was performed using DISTRUCT v.1.1 (Rosenberg, 2004). Based on the same data set, a principal component analysis (PCA) was performed with ten principal components (PCs) extracted in PLINK v.1.90 (Purcell et al., 2007), and visualized using R.

2.4. Phylogenetic inference and divergence time estimation

First, a nuclear phylogenetic tree based on SNPs was generated using maximum likelihood (ML) in IQ-TREE v.1.6.8 (Nguyen et al., 2015) with 1000 ultrafast bootstraps (Hoang et al., 2018). The Python script 'vcf2phylip' (<https://github.com/edgardomortiz/vcf2phylip>) was used to transfer the SNP data for tree construction. *Gentiana crassuloides* (specimen no. Fu2019013) was used as an outgroup.

Subsequently, the assembled plastome sequences were used to construct an ML tree. Following the exclusion of one inverted repeat (IR) region, the plastome sequences were aligned using MAFFT (Katoh et al., 2002). The most rapidly evolving sites were eliminated using Gblocks (Talavera and Castresana, 2007) with default settings. The substitution model was selected using ModelFinder (Kalyaanamoorthy et al., 2017). Phylogenetic analyses were performed using IQ-TREE (Nguyen et al., 2015) based on the ML and with 5000 ultrafast bootstrap replicates (Hoang et al., 2018).

The divergence times of the main lineages of *Gentiana aristata* were estimated based on the plastome (removed one IR) matrix using the Bayesian method implemented in BEAST 2.4 (Drummond et al., 2012; Bouckaert et al., 2014). The analyses were performed using the HKY substitution, the Yule, and the strict clock model. Because no *G. aristata* fossils have been discovered, the dating was calibrated in two ways to increase its accuracy. First, the stem node of *G. section Cruciata* was constrained with a section of a *Cruciata* fossil from the early Miocene (Mai, 2000). Lognormal priors with an offset at 16.0 million years ago (Ma), a mean of 1, and a standard deviation of 1.0 were used. As a result of very limited fossil evidence for gentians, the crown age of *Gentiana* was also constrained after including primary lineages of this genus. Uniform priors (Schenk

and Axel, 2016) with a lower age of 21 Ma and an upper age of 38 Ma were used to integrate the entire 95% Highest Posterior Density (HPD) from Janssens et al. (2020). An analysis of priors solely (without sequence data) was performed to determine if there was prior interaction (Warnock et al., 2015). Three independent Monte Carlo Markov chains (MCMC) with 10 million generations, sampling every 1000th generation, were run with the initial 20% discarded as burn-in. Effective Sample Size (ESS) values (> 200) were utilized to determine whether the convergence was appropriate, and TreeAnnotator 1.7.5 was used to summarize the tree (Drummond et al., 2012). Information on outgroup sequences is presented in Table S1.

2.5. Hybridization analysis

Patterson's D -statistics (Durand et al., 2011) were employed to assess introgression within *Gentiana aristata*. The D -statistics use asymmetry in gene tree topologies to quantify introgression between either of two lineages that share a common ancestor (P1 and P2) and one additional lineage (P3) that diverged from the common ancestor of P1 and P2 earlier. Patterson's D -statistics and f_4 -ratio for all possible population trios were determined using the *Dtrios* function of Dsuite v.0.5.r44 (Malinsky et al., 2021) with default parameters. *G. crassuloides* was fixed as the outgroup. The significance of each test was evaluated using 100 jackknife resampling runs. Results from *Dtrios* were further processed using the f -branch program (Malinsky et al., 2021) and its associated plotting utilities for the f -branch statistics (Malinsky et al., 2018).

2.6. Species distribution modeling

Current distribution data for *Gentiana aristata* were obtained through our fieldwork records and the Global Biodiversity Information Facility (GBIF.org (03 March 2023) GBIF Occurrence Download <https://doi.org/10.15468/dLx39qqb>). Records occurring less than 10 km from each other were eliminated in ArcGIS 10.2 to prevent multicollinearity. The 19 bioclimatic variables utilized at present, Last Glacial Maximum (LGM, ~22 kya), and Last Interglacial (LIG, ~120–140 kya) were obtained from the WorldClim2 data set (Fick and Hijmans, 2017) with a spatial resolution of 2.5 arc-min (30 arc-sec for LIG). Highly correlated variables were eliminated according to Pearson's correlation coefficient, which was significantly higher than 0.9 (p -value < 0.05). Using MaxEnt 3.4.1 (Phillips et al., 2017), the potential distribution area was predicted. Approximately 75% of the location data was used for training, while the remaining 25% was used to test the predictive ability of the model. The effectiveness of the model was assessed using the receiver operating characteristic (ROC) and the area under the ROC curve (AUC) > 0.9.

3. Results

3.1. Genetic diversity and divergence

Illumina sequencing of GBS libraries generated an average of 3,324,959 clean reads and 355,828 tags per sample. A total of 1,218,202 SNPs were identified from all samples against the genome, and 29,865 unlinked SNPs were left for downstream analysis following filtering for MAF, LD, and missing data. In *Gentiana aristata*, the mean A_r , H_0 , and F_{IS} were 1.199, 0.125, and 0.417, respectively. Genomic SNPs revealed that two western populations (YS and LWQ) had a slightly lower genetic diversity (e.g., A_r , H_0) in comparison to the remaining populations (Table 1). The populations YS and LWQ also had higher values of pairwise F_{ST} (0.436–0.529) in comparison to the remaining populations

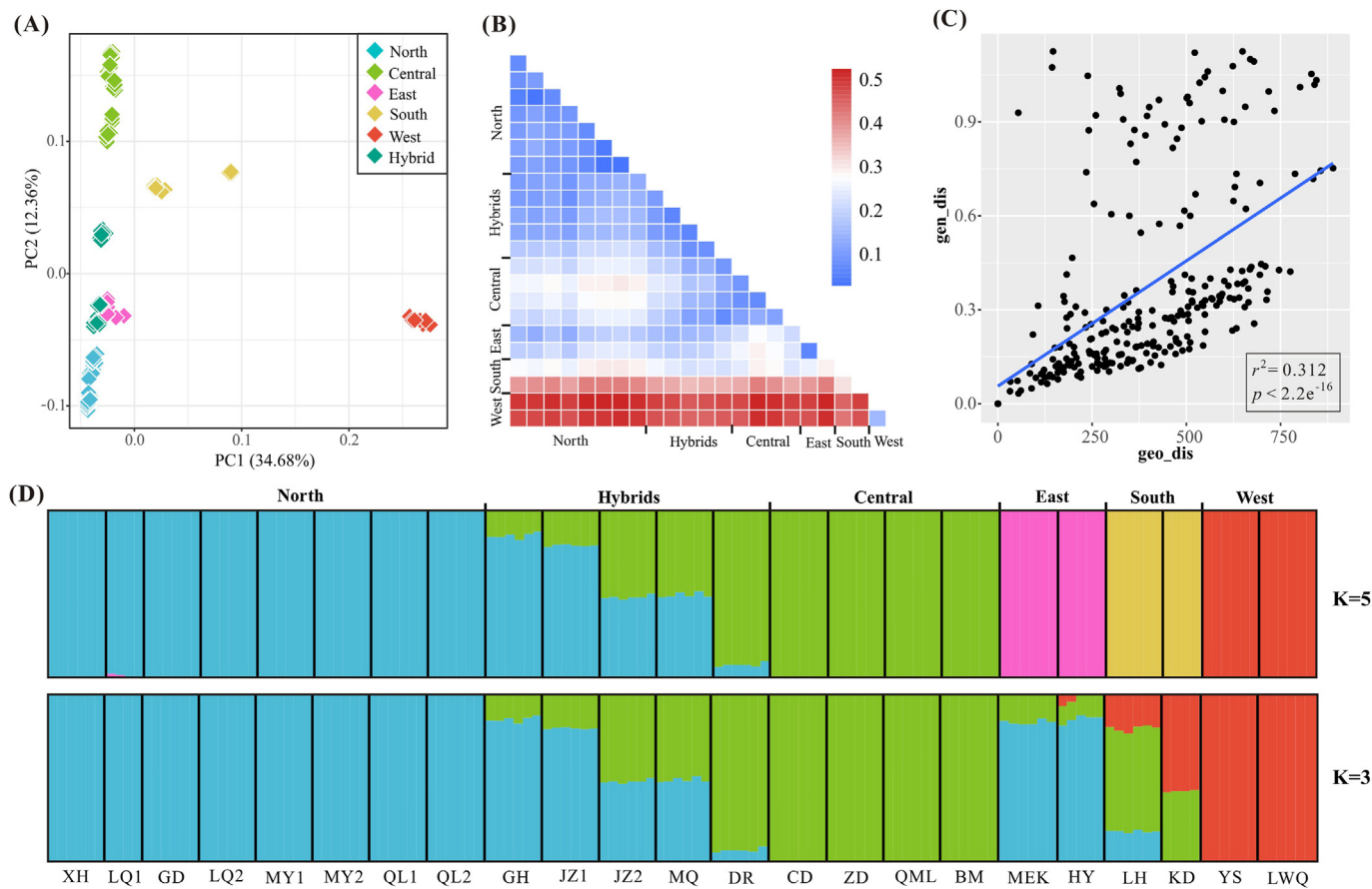


Fig. 2. Genetic clustering and differentiation in *Gentiana aristata*. (A) Results of principal coordinate analysis. Each square represents an individual. (B) A heatmap of the weighted Weir and Cockerham's F_{ST} . (C) Genetic distance ($F_{ST}/(1-F_{ST})$) versus geographical distance (km) between populations. (D) Bar plots indicating probabilities of ancestral clusters of each sample at $K = 5$ and 3 in FastStructure.

(0.040–0.429) (Fig. 2B and Table S2). The plot of $F_{ST}/(1-F_{ST})$ versus geographic distances among populations revealed a significant positive correlation ($r^2 = 0.312$, $p < 2.2e^{-16}$), indicating strong signals of IBD (Fig. 2C).

3.2. Population genetic structure and hybridization

The marginal likelihood scores from FastSTRUCTURE analyses peaked at $K = 5$ (Fig. S1), indicating the populations can be divided into five genetic groups (the West, East, Central, South and North). In general, the inferred genetic groups corresponded with known geographic barriers present between populations. The North cluster (comprising populations QL1, QL2, MY1, MY2, GD, XH, LQ1, and LQ2) stretches from the Qilian Mountains to Min Mountains (Figs. 1 and 2D). The Central cluster (ZD, QML, CD, and BM) is mainly situated between the Bayankala Mountains and the Yangtze River, while the West cluster is located to the west of the Yangtze River. The exception is the population ZD, which was grouped into the Central cluster but is located to the west of the Yangtze River (Figs. 1 and 2D). The East and South clusters are both located in the Hengduan Mountains. Five populations (GH, MQ, JZ2, JZ1, and DR) situated along the Bayankala Mountains were named Hybrids as a result of their combination of North and Central clusters. The proportion of the North cluster in populations GH, JZ1, MQ, JZ2 and DR, was 0.85, 0.79, 0.49, 0.48, and 0.08, respectively. The marginal likelihood scores from FastSTRUCTURE did not sharply increase since $K = 3$ (Fig. S1). Genetic grouping at $K = 3$ showed that the East and South were mixed by a North, Central or West cluster (Fig. 2D).

The PCA plot identified six clusters, corresponding to the North, Central, West, East, South and Hybrids (Fig. 2A). The first two principal components, PC1 and PC2, explained 34.68% and 12.36% of the total variation, respectively. PC3 explained 8.05% of total variation and differentiated the East and Hybrid clusters (Fig. S2).

The f -branch statistics revealed considerable gene flow from the Central cluster to Hybrids and the Hybrids to the North, as well as the West cluster to the base of the other four clusters. At the population level, strong gene flow was detected from CD, BM and DR to MQ and JZ2, and evident gene flow from JZ2 to JZ1 (Fig. 3). Evident gene flow was observed between almost all clusters except West and South to DR, which was located in the center of the range (Fig. 3).

3.3. Phylogenetic inference and estimation of divergence time

The genomic SNPs generated a well-supported ML tree in which all nodes were fully supported (100%) except for three nodes with support values ranging from 89% to 99% (Fig. 4). Individuals from the same population were clustered into one subclade, except for those of HY and MEK, which were mixed in one subclade. Populations from the West (YS and LWQ) were located at the base of the tree, and were sister groups with the remaining populations. Populations from the Hybrids group were located between the Central and the North, the latter of which were assigned to a clade at the tip of the tree (Fig. 4).

Using the genome-skimming data, 23 plastomes were assembled with lengths ranging between 127,562 bp and 128,190 bp.

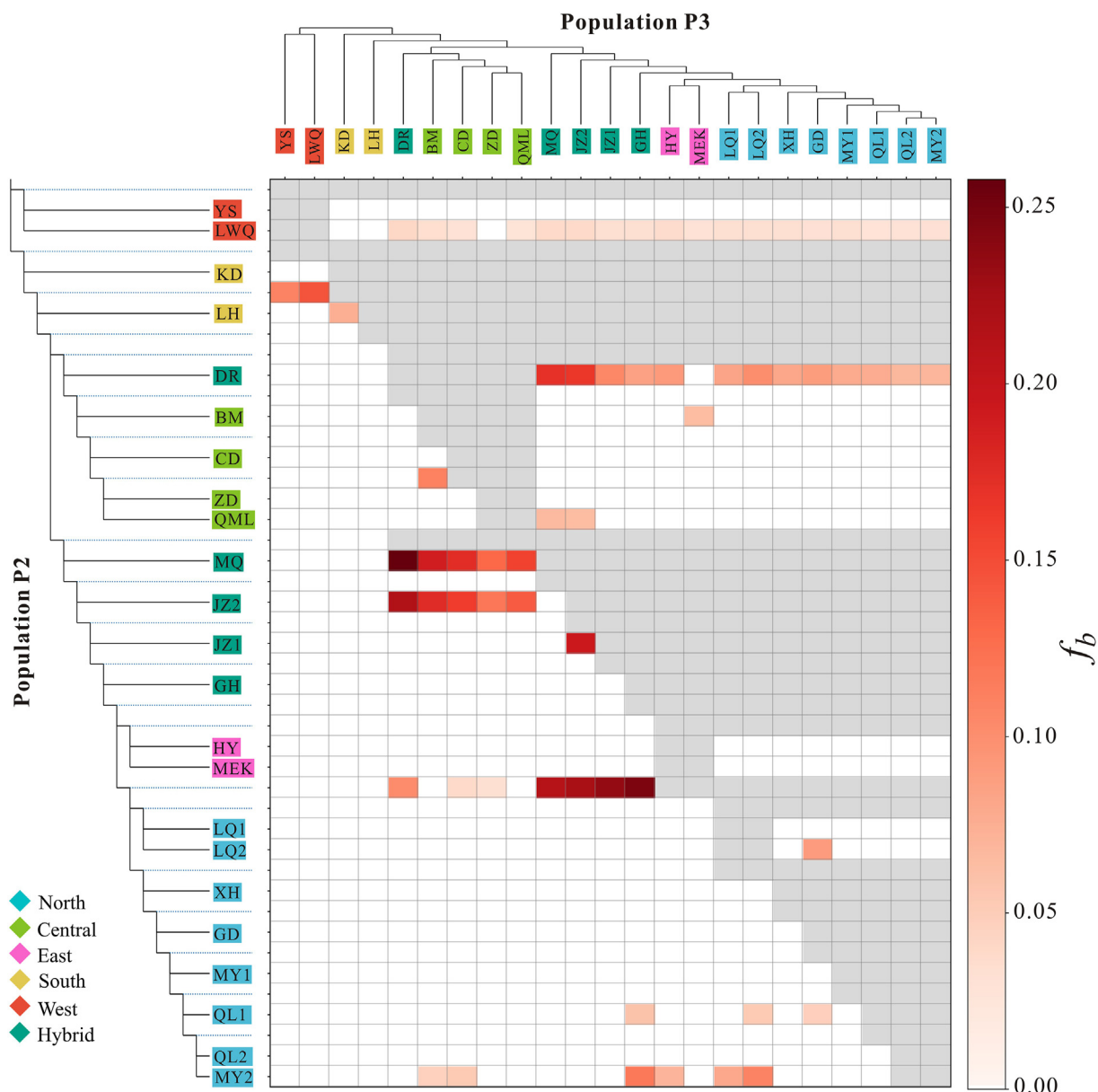


Fig. 3. Gene flow detected in *Gentiana aristata*. A heatmap summarizing the f_b -branch (f_b) statistics from Dsuite. The population tree based on nuclear SNPs is displayed along the sides. The values in the matrix refer to excess allele sharing between the branch b identified on the extended tree on the y-axis and the population P3 identified on the x-axis. The color of each square denotes the f_b -branch statistics estimate. The colors in the population names represent the genetic clusters in Fig. 2.

After one IR was eliminated, the plastome matrix with sequence lengths from 105,342 bp to 105,853 bp was used for downstream analysis. The plastome matrix generated a robust backbone in the ML tree (Fig. 4). The West as well as populations CD and ZD from the Central cluster were located at the base of the tree. The trees based on SNP and plastome data revealed differences in the placement of most populations, indicating substantial cyto-nuclear conflicts. For example, population LH from the South clustered together with populations BM and QML from the Central in the plastome tree. Populations belonging to Hybrids were assigned to distinct clades in the plastome tree (Fig. 4).

Our divergence time analyses based on the plastome matrix indicated that the population YS and the remaining populations diverged in the Early Pliocene, approximately 4.24 Ma (95% HPD: 3.95–4.52 Ma). The remaining three clades diverged during the

Early Quaternary, approximately between 1.10 and 1.46 Ma (Fig. 5). The detailed results, including for the outgroup, are illustrated in Fig. S3.

3.4. Palaeo-distributional reconstruction

A total of 61 filtered occurrences were used for species distribution modeling analysis. Following Pearson’s correlation analysis, ten bioclimatic variables (bio2–bio5, bio7, bio9, bio13–bio15, and bio18) were used for distributional reconstruction. The palaeo-distributional reconstruction showed that two small disjunctive areas in the Hengduan Mountains served as the potential habitat of *Gentiana aristata* during the LIG (Fig. 6). The potential habitat of *G. aristata* increased northward in its eastern range during the LGM, and then further expanded northward until present (Fig. 6).

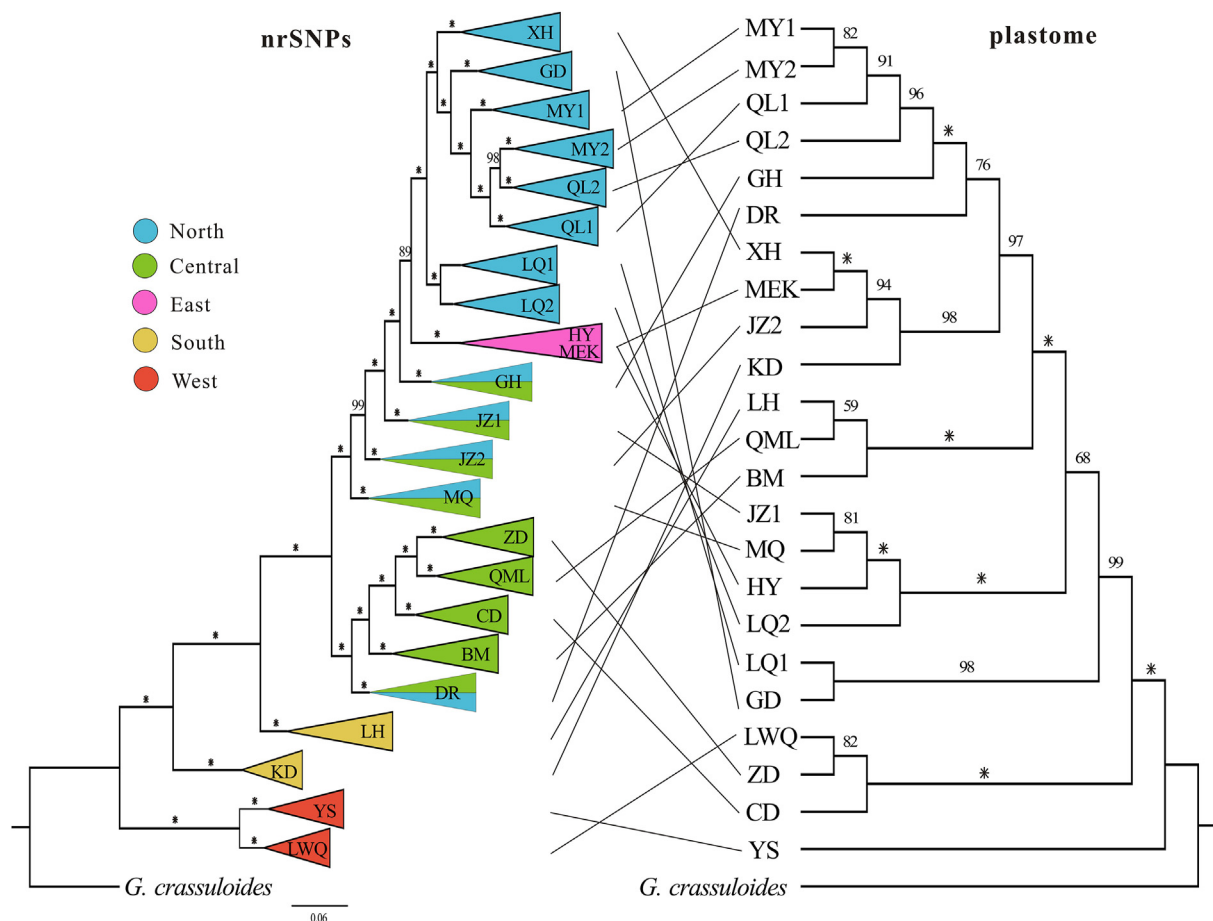


Fig. 4. Maximum likelihood trees of *Gentiana aristata* based on nuclear SNPs and plastome sequences. Phylogenetic support values are shown above branches only when they deviate from 100% bootstrap support, which is shown as an asterisk (*). Codes in the tips indicate the population names. Each population included in the phylogenetic trees is connected with a line. The colors represent the genetic clusters in Fig. 2.

4. Discussion

As a region with dynamic diversification, secondary contact and speciation, the QTP region offers an excellent opportunity to analyze the evolutionary history of diversification and how it has been influenced by geological or climatic modifications. Focusing on an annual species in the most speciose lineage of *Gentiana*, we observed deep genetic divergence corresponding to geological barriers in the QTP region and recovered cryptic diversity based on genetic and morphological evidence. In populations along the Bayankala Mountains, hybridization was rampant. Here, we discuss our results in relation to the biogeographic history of this significant hotspot and contemplate the implications for future studies.

4.1. Strong genetic differentiation and cryptic diversity in *Gentiana aristata*

The QTP region is renowned for its extraordinary diversity and high rate of *in situ* alpine speciation (Sun et al., 2017; Xing and Ree, 2017; Ding et al., 2020). However, the way that genetic subdivision of populations and biodiversity assemble in speciose lineages within this region has yet to be characterized with precision. Our study addresses this issue by using genomic data and fine-scale population to reveal deep genetic differentiation in *Gentiana aristata*. Our genetic clustering disentangled five distinct clusters in *G. aristata*, and observed great genetic differentiation between the West and the remaining clusters (average $F_{ST} = 0.491$). Indeed,

differentiations within species having wide HM ranges are commonly significant, e.g., *Paraquilegia microphylla* ($F_{ST} = 0.919–0.926$; Luo et al., 2016), *Meconopsis integrifolia* ($F_{ST} = 0.726–0.905$, Yang et al., 2012), *Circaea agrestis* ($F_{ST} = 0.128–0.891$; Zhang et al., 2020), and *Gentiana szechenyii* ($F_{ST} = 0.762–0.763$; Sun et al., 2022). The extreme ruggedness and geographical barriers in the QTP region, in particular in the HM, are significant factors for the considerable differentiation in diverse species in this area. We also observed a high level of average inbreeding coefficient in *G. aristata* ($F_{IS} = 0.417$), which could limit gene flow and increase differentiation among populations. The character of herkogamy in *G. aristata* indicates that selfing in this species shall be rare, but tiny pollinators such as ants could indeed cause selfing, as for example in *G. lawrencei* (Hou et al., 2009).

Geographical features, such as mountains and rivers, could act as barriers to gene flow, leading to genetic differentiation and even speciation (Seregin et al., 2015; Pertierra et al., 2020; White et al., 2020). The five distinct clusters occur in five areas (Fig. 1) that are geographically separated by physical barriers, such as Bayankala Mountains and the Yangtze River. The Bayankala Mountains, often referred to as Bayan Har Mountains, are a spur of the Kunlun Mountains in China and serve as the watershed for the headwaters of the Yellow River and Yangtze River. Although the Bayankala Mountains are not characterized by high ruggedness as the HM, their average elevation of approximately 5000 m makes them sufficient to serve as barriers throughout glacial cycles for species like *Gentiana aristata* with short-distance dispersal pollinators.

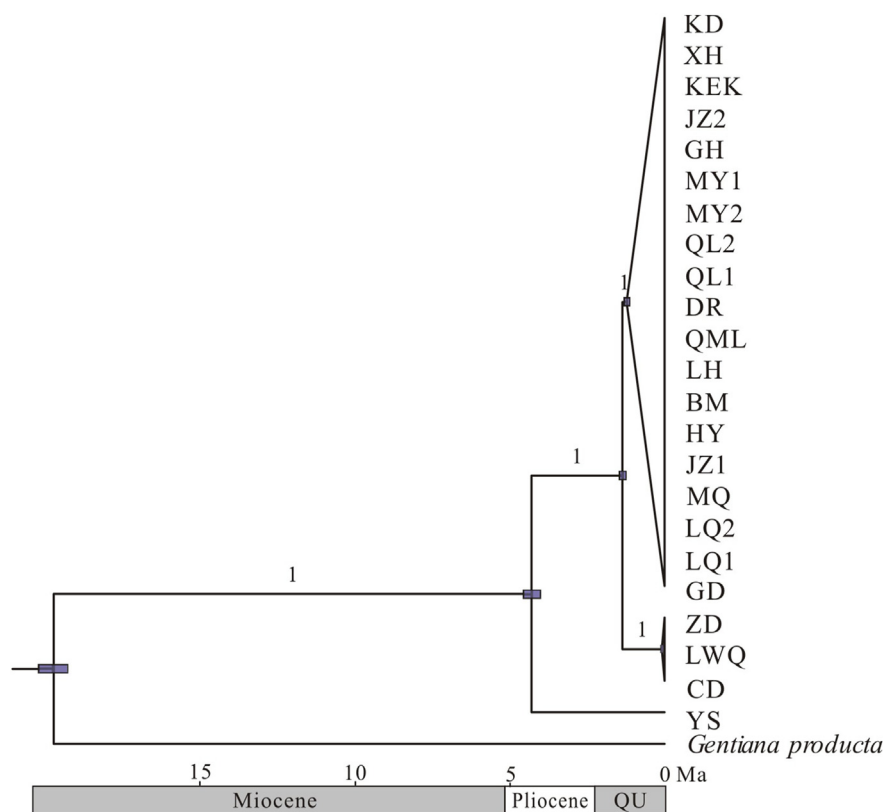


Fig. 5. Divergence time estimation in *Gentiana aristata*. The gray bars indicate the 95% highest posterior density on the age estimates. The posterior probabilities of nodes are displayed above branches. Ma, million years ago; QU, the Quaternary.

Recently, a set of studies on plants from the QTP region, including *Pinus armandii* (Liu et al., 2019), *Rheum palmatum* (Feng et al., 2020), and *Gentiana hexaphylla* (Fu et al., 2022b), identified a distinct genetic pattern in this topographically complicated environment. These geographical characteristics also had an important impact on species assembly across the HM and the greater QTP region (Li et al., 2021).

Climate changes associated with glacial cycles may also significantly alter the distribution of species. Specifically, climate changes may fragment or shift the ranges of species, accelerating diversification or secondary contact within montane species (Muellner-Riehl, 2019), including *Gentiana* (Fu et al., 2022a). The divergence time within *G. aristata* primarily occurred in the Early Quaternary when the QTP region underwent cycles of glaciations and plant species experienced recurrent range changes to track favorable habitats (Hewitt, 2000; Zheng et al., 2002). Studies of *Anisodus tanguticus* (Wan et al., 2016), *Medicago* (Guo et al., 2023), and *Ostryopsis intermedia* (Liu et al., 2014), as well as those on birds (Lei et al., 2014), also suggest lineage diversification during the Early Quaternary in this region. Therefore, the existing genetic pattern in *G. aristata* may have been fostered by both geographical features and climate oscillations.

In areas comprising rich and dynamic flora like the QTP region, morphological variation is frequently observed within species of diverse lineages, including *Gentiana* (Ho and Pringle, 1995; Ho and Liu, 2001). The flora described *G. aristata* as having blue to pale purple corollas (Ho and Pringle, 1995), which is consistent with our observations that the corollas are blue in populations YS and LWQ or pale purple in the remaining populations (Fig. 1). Although corolla color has revealed phenotypic plasticity in a number of species (Clegg and Durbin, 2000; Muchhala et al., 2014;

Liang et al., 2023), populations with blue corollas in this study exhibited deep genetic differentiation from those with pale purple corollas, as indicated by the high F_{ST} values (0.436–0.529), distinct genetic clustering (Fig. 2A and D), and sister placement in the phylogenetic trees (Fig. 4). Additionally, populations in the West cluster (YS and LWQ) are geologically separated from the remaining clusters by the Tongtian River, which is the upper reaches of the Yangtze River. The divergence period between population YS and the remaining populations occurred in the Early Pliocene, when the Tongtian River already existed (Chen et al., 2020; Zhang et al., 2021). Meanwhile, the divergence time between population YS and the remaining populations occurred as early as approximately 4.15 Ma, a time range ensuring divergence occurred between morphologically distinct species in gentian (Fig. S3; Favre et al., 2016). Interestingly, populations in the West cluster revealed slightly lower genetic diversity in comparison to the remaining populations, considering high F_{IS} values in *G. aristata*, together exhibiting poor gene flow between the West and the remaining populations, which was supported by D -statistics (Fig. 3). Taken together, these findings suggest that the cryptic diversity in populations YS and LWQ were caused by geological isolation. Population ZD, located west of the Yangtze River, was grouped into the Central cluster rather than the West. This clustering might be explained by weak geological isolation near the origin of the Yangtze River. Genetic data, especially genomic data, have provided growing evidence of cryptic species in both animals and plants (Struck et al., 2018), for example, in butterflies (Hebert et al., 2004), bumblebees (Christmas et al., 2021), yew (Liu et al., 2013), and *Sarracenia* (Carstens and Satler, 2013). Thus, we propose that the cryptic diversity in dynamic flora likely in the QTP region might be vastly underestimated, and

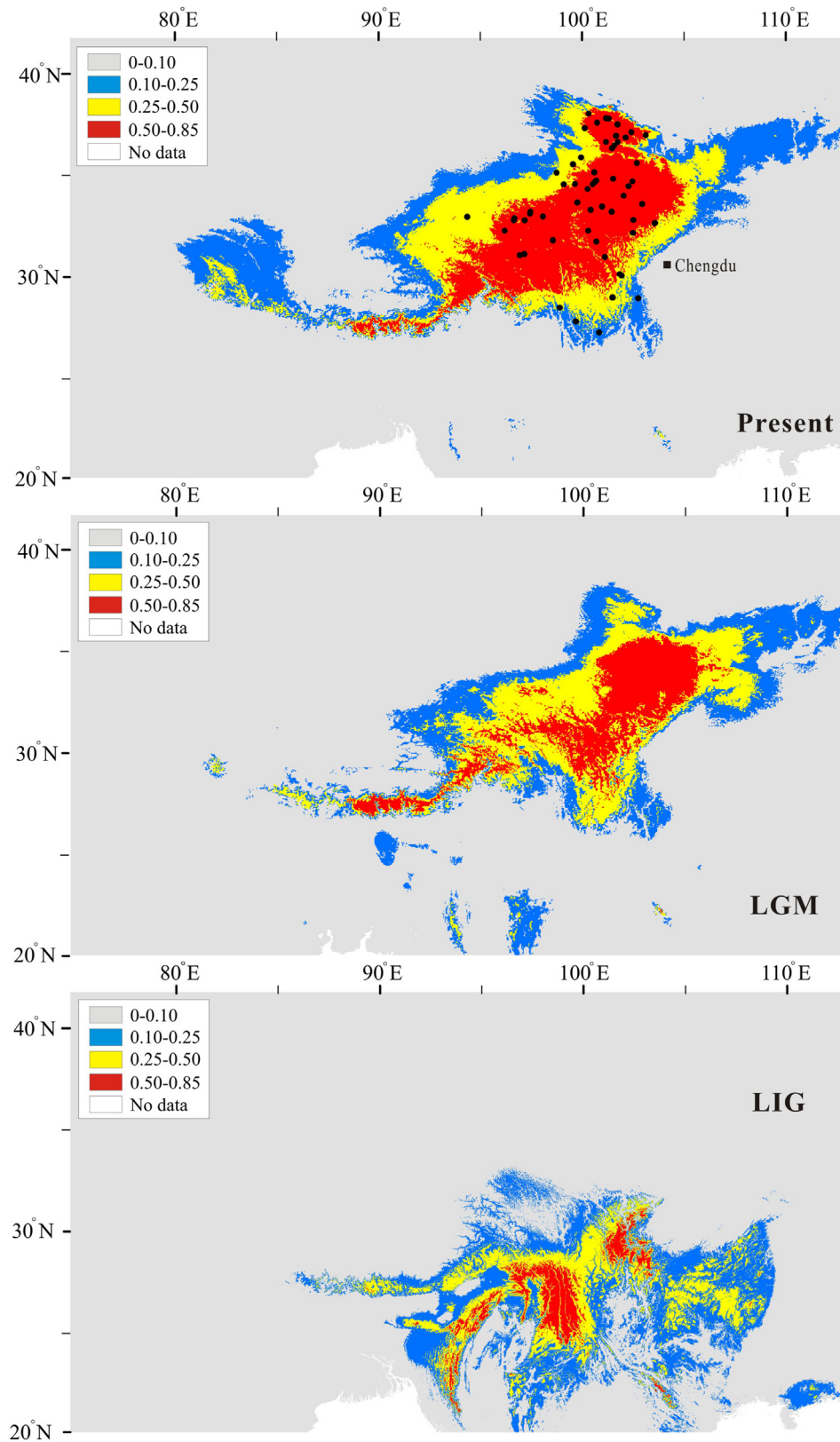


Fig. 6. Estimated climatic niches for distribution of *Gentiana aristata* in the Qinghai-Tibet Plateau region. Maps are shown at present, LGM (~22 kya) and LIG (~120–140 kya). The value of predicted habitat suitability is indicated by the bars in each panel. Current distribution records are represented in the present map by black dots.

genomic data should be widely employed to evaluate the biodiversity of the region.

4.2. Rampant hybridization along Bayankala Mountains

The QTP region offers an excellent opportunity to gather evidence on the role of hybridization and introgression for adaptation and diversification because of the remarkable number of radiations in the QTP region, of which many closely related species are sympatric. Our genomic SNPs not only disentangled distinct clusters in allopatry in *Gentiana aristata* but also identified a very clear intra-species hybrid region around the Bayankala Mountains (Fig. 1). Regarding the structure of the hybrid region, the proportion of the North cluster, for example, in the five sampled populations ranged from 0.08 to 0.85, indicating that it consisted of offspring of recent hybrids or backcrosses. Interestingly, phylogenetic analyses based on plastid data demonstrated that the five populations are clustered into various lineages (Fig. 4), indicating the different maternal origins of the hybrids and bi-directional gene flow in this hybrid region. Moreover, genetic grouping at $K = 3$ and PCA plots both indicated that the East and South clusters are mixed with the North, Central or West clusters (Fig. 2A and D), indicating they are likely introgressed rather than divergent lineages, as for example in white birch (Nocchi et al., 2023), although more evidence is needed. In the East and South clusters located in the HM, which is characterized by deeply dissected landscapes and has created complex genetic structure, geological isolation and hybridization are likely the main factors that have shaped genetic differentiation. In fact, hybridization has been widely observed to drive species formation and biodiversity assembly in the QTP region (Wu et al., 2022), from herbs (e.g., *Saxifraga*, Ebersbach et al., 2020) and shrubs (e.g., *Rhododendron*, Ma et al., 2022) to woody species (e.g., *Pterocarya*, Wang et al., 2023; *Populus*, Li et al., 2023).

Because the Bayankala Mountains separate populations into two genetic clusters, any observed hybridization here would likely be caused by secondary contact following range changes induced by climate oscillations. Climate oscillations cause variable connectivity across sky-islands in mountain systems as demonstrated, for example, in the HM (Deng et al., 2020) and Flickering Connectivity System (Andes; Flantua et al., 2019; Dantas-Queiroz et al., 2023). In a number of cases, connectivity in the QTP region has been shown to facilitate hybridization across different genetic groups at intra- and inter-species levels (Liu et al., 2014; Wu et al., 2022). Although the Bayankala Mountains served as one of the major geological barriers to populations of *Gentiana aristata*, vertical displacement as a result of temperature changes have offered chances for clusters on opposite sides of the mountain to come into contact in some areas. Our species distribution modeling revealed range expansion since LIG in *G. aristata* (Fig. 6), as expected. In particular, the potential range size of most alpine plants expanded, rather than contracted, during the warming in the QTP region (Liang et al., 2018), providing additional opportunities for contact in this region in the context of climate warming.

Natural hybrids can be identified in almost all species-rich genera in the QTP region (Yang et al., 2019; Wu et al., 2022). Growing evidence from both population genetics and phylogenetics suggests that hybridization might be more pervasive than previously believed in this species-rich genus *Gentiana*, which has received increased attention. Hybridization was observed in nearly every taxon examined, for example, ancient hybridization in the sections *Ciminalis* (Christe et al., 2014), *Cruciata* (Fu et al., 2021a), and *Chondrophyllae* (Chen et al., 2021; Fu et al., 2022c), and recent hybridization in the sections *Cruciata* (Hu et al., 2016), *Kudoa* (Fu et al., 2020, 2022a), and *Isomeria* (Sun et al., 2022). Compared with rapidly growing cases of hybridization within the genus,

reports of ancient hybridization among genera or lineages are scarce. Recent studies of phylotranscriptomics in the subtribe Swertiinae (Gentianaceae) demonstrate that a large clade comprising 10 genera has an ancient hybrid origin (Chen et al., 2023), suggesting hybridization might be more frequent than previously assumed and could be technically identified in the genomic era. The role of hybridization in rapid diversification events and biodiversity assembly in dynamic flora, such as in the QTP region and Andean páramos, are still relatively understudied despite growing evidence (Schley et al., 2022) and increasing genomic resources that could help shed light on the evolutionary effects of hybridization.

5. Conclusions

In this study, we used genomic data to reveal deep differentiation of *Gentiana aristata* in the QTP region, and rampant hybridization resulting in post-F1 hybrids along a geographical feature. We identified cryptic diversity in *G. aristata* with evidence obtained from both genetic and morphological characteristics. The differentiation and cryptic diversity were promoted by climatic fluctuations and geological barriers. These findings suggest that cryptic diversity and hybridization are crucial for biodiversity assembly in the QTP region.

Author contributions

P.C.F. designed the experiments, collected the samples, analyzed the data and wrote the main manuscript; Q.Q.G. and D.C. analyzed the data and prepared figures; Q.B.G. collected the samples; S.S.S. analyzed the data and revised the manuscript. All the authors have read and approved the final manuscript.

Data availability statement

The data set supporting the conclusions of this article is available in the Figshare (<https://doi.org/10.6084/m9.figshare.22698871>).

Declaration of competing interest

The authors have no conflict of interest.

Acknowledgements

We thank Yi-Lin Fan of Zhengzhou Customs Technology Center for the support in sample collecting. We are grateful to financial support provided by the Foundation of Henan Educational Committee (22A180024) and Natural Science Foundation of Henan Province (232300420212).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pld.2023.10.004>.

References

- Abbott, R.J., 2017. Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *J. Syst. Evol.* 55, 238–258.
- Abbott, R.J., Albach, D., Ansell, S., et al., 2013. Hybridization and speciation. *J. Evolution. Biol.* 26, 229–246.
- Antonelli, A., Kissling, W.D., Flantua, et al., 2018. Geological and climatic influences on mountain biodiversity. *Nat. Geosci.* 11, 718–725.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.

- Bouckaert, R., Heled, J., Kühnert, D., et al., 2014. Beast 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10, e1003537.
- Carstens, B.C., Satler, J.D., 2013. The carnivorous plant described as *Sarracenia alata* contains two cryptic species. *Biol. J. Linn. Soc.* 109, 737–746.
- Chen, J., 2020. Evolutionary Process of the Yangtze River. *Evolution and Water Resources Utilization of the Yangtze River*. Springer, Singapore, pp. 47–122.
- Chen, C.L., Zhang, L., Li, J.L., et al., 2021. Phylotranscriptomics reveals extensive gene duplication in the subtribe Gentianinae (Gentianaceae). *J. Syst. Evol.* 59, 1198–1208.
- Chen, C., Ruhfel, B.R., Li, J., et al., 2023. Phylotranscriptomics of Swertiae (Gentianaceae) reveals that key floral traits are not phylogenetically correlated. *J. Integr. Plant Biol.* 65, 1490–1504.
- Christe, C., Caetano, S., Aeschimann, D., et al., 2014. The intraspecific genetic variability of siliceous and calcareous *Gentiana* species is shaped by contrasting demographic and re-colonization processes. *Mol. Phylogenet. Evol.* 70, 323–336.
- Christmas, M.J., Jones, J.C., Olsson, A., et al., 2021. Genetic barriers to historical gene flow between cryptic species of alpine bumblebees revealed by comparative population genomics. *Mol. Biol. Evol.* 38, 3126–3143.
- Clegg, M.T., Durbin, M.L., 2000. Flower color variation: a model for the experimental study of evolution. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7016–7023.
- Danecek, P., Auton, A., Abecasis, G., et al., 2011. The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158.
- Dantas-Queiroz, M.V., Hurbath, F., de Russo Godoy, F., et al., 2023. Comparative phylogeography reveals the demographic patterns of Neotropical ancient mountain species. *Mol. Ecol.* 32, 3165–3181.
- Deng, J.Y., Fu, R.H., Compton, S.G., et al., 2020. Sky islands as foci for divergence of fig trees and their pollinators in southwest China. *Mol. Ecol.* 29, 762–782.
- Ding, W.N., Ree, R.H., Spicer, R.A., et al., 2020. Ancient orogenic and monsoon-driven assembly of the world's richest temperate alpine flora. *Science* 369, 578–581.
- Drummond, A.J., Suchard, M.A., Xie, D., et al., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973.
- Durand, E.Y., Patterson, N., Reich, D., et al., 2011. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252.
- Ebersbach, J., Tkach, N., Röser, M., et al., 2020. The role of hybridisation in the making of the species-rich arctic-alpine genus *Saxifraga* (Saxifragaceae). *Diversity* 12, 440.
- Elsen, P.R., Tingley, M.W., 2015. Global mountain topography and the fate of montane species under climate change. *Nat. Clim. Change* 5, 772–776.
- Favre, A., Päckert, M., Pauls, S.U., et al., 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biol. Rev.* 90, 236–253.
- Favre, A., Michalak, I., Chen, C.H., et al., 2016. Out-of-Tibet: the spatio-temporal evolution of *Gentiana* (Gentianaceae). *J. Biogeogr.* 43, 1967–1978.
- Favre, A., Pringle, J.S., Heckenhauer, J., et al., 2020. Phylogenetic relationships and sectional delineation within *Gentiana* (Gentianaceae). *Taxon* 69, 1221–1238.
- Favre, A., Paule, J., Ebersbach, J., 2021. Incongruences between nuclear and plastid phylogenies challenge the identification of correlates of diversification in *Gentiana* in the European alpine system. *Alp. Bot.* 132, 29–50.
- Feng, L., Ruhsam, M., Wang, Y.H., et al., 2020. Using demographic model selection to untangle allopatric divergence and diversification mechanisms in the *Rheum palmatum* complex in the Eastern Asiatic region. *Mol. Ecol.* 29, 1791–1805.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315.
- Flantua, S.G., O'Dea, A., Onstein, R.E., et al., 2019. The flickering connectivity system of the north Andean páramos. *J. Biogeogr.* 46, 1808–1825.
- Fu, P.C., Ya, H.Y., Liu, Q.W., et al., 2018. Out of refugia: population genetic structure and evolutionary history of the alpine medicinal plant *Gentiana lawrencei* var. *farreri* (Gentianaceae). *Front. Genet.* 9, 564.
- Fu, P.C., Sun, S.S., Khan, G., et al., 2020. Population subdivision and hybridization in a species complex of *Gentiana* in the Qinghai-Tibetan Plateau. *Ann. Bot.* 125, 677–690.
- Fu, P.C., Twyford, A.D., Sun, S.S., et al., 2021a. Recurrent hybridization underlies the evolution of novelty in *Gentiana* (Gentianaceae) in the Qinghai-Tibetan Plateau. *AoB Plants* 13, plaa068.
- Fu, P.C., Sun, S.S., Twyford, A.D., et al., 2021b. Lineage-specific plastid degradation in subtribe Gentianinae (Gentianaceae). *Ecol. Evol.* 11, 3286–3299.
- Fu, P.C., Sun, S.S., Hollingsworth, P.M., et al., 2022a. Population genomics reveal deep divergence and strong geographical structure in gentians in the Hengduan Mountains. *Front. Plant Sci.* 13, 936761.
- Fu, P.C., Favre, A., Wang, R., et al., 2022b. Between allopatry and secondary contact: differentiation and hybridization among three sympatric *Gentiana* species in the Qinghai-Tibet Plateau. *BMC Plant Biol.* 22, 1–15.
- Fu, P.C., Chen, S.L., Sun, S.S., et al., 2022c. Strong plastid degradation is consistent within section *Chondrophyllae*, the most speciose lineage of *Gentiana*. *Ecol. Evol.* 12, e9205.
- Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. *arXiv*, 1207.3907.
- Goudet, J., 2005. HIERFSTAT, a package for r to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* 5, 184–186.
- Guo, W., Yang, Y., Zhang, X., et al., 2023. Genomic divergence between two sister *Medicago* species triggered by the Quaternary climatic oscillations in the Qinghai-Tibet Plateau and northern China. *Mol. Ecol.* 32, 3118–3132.
- Hebert, P.D., Penton, E.H., Burns, J.M., et al., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14812–14817.
- Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405, 907–913.
- Ho, T.N., Liu, S.W., 2001. A Worldwide Monograph of *Gentiana*. Science Press, Beijing.
- Ho, T.N., Pringle, J.S., 1995. Gentianaceae. In: Wu, Z.Y., Raven, P.H. (Eds.), *Flora of China*, vol. 16. Science Press, Beijing & Missouri Botanical Garden, St. Louis, pp. 1–140.
- Hoang, D.T., Chernomor, O., Von Haeseler, A., et al., 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522.
- Hou, Q.Z., Duan, Y.W., Si, Q.W., et al., 2009. Pollination ecology of *Gentiana lawrencei* var. *farreri*, a late-flowering Qinghai-Tibet plateau species. *Chin. J. Plant Ecol.* 33, 1156–1164.
- Hu, Q., Peng, H., Bi, H., et al., 2016. Genetic homogenization of the nuclear ITS loci across two morphologically distinct gentians in their overlapping distributions in the Qinghai-Tibet Plateau. *Sci. Rep.* 6, 1–11.
- Janssens, S.B., Couvreur, T.L.P., Mertens, A., et al., 2020. A large-scale species level dated angiosperm phylogeny for evolutionary and ecological analyses. *BioRxiv*. Data J. 8, e39677.
- Jin, J.J., Yu, W.B., Yang, J.B., et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21, 1–31.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., et al., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587.
- Katoh, K., Misawa, K., Kuma, K.I., et al., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Kirschner, P., Závieská, E., Gamisch, A., et al., 2020. Long-term isolation of European steppe outposts boosts the biome's conservation value. *Nat. Commun.* 11, 1968.
- Lei, F., Qu, Y., Song, G., 2014. Species diversification and phylogeographical patterns of birds in response to the uplift of the Qinghai-Tibet Plateau and Quaternary glaciations. *Curr. Zool.* 60, 149–161.
- Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., et al., 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, Q., Sun, H., Boufford, D.E., et al., 2021. Grade of membership models reveal geographical and environmental correlates of floristic structure in a temperate biodiversity hotspot. *New Phytol.* 232, 1424–1435.
- Li, T., Yu, X., Ren, Y., et al., 2022. The chromosome-level genome assembly of *Gentiana dahurica* (Gentianaceae) provides insights into gentiopicroside biosynthesis. *DNA Res.* 29, dsac008.
- Li, X., Ruhsam, M., Wang, Y., et al., 2023. Wind-dispersed seeds blur phylogeographic breaks: the complex evolutionary history of *Populus lasiocarpa* around the Sichuan Basin. *Plant Divers.* 45, 156–168.
- Liang, Q., Xu, X., Mao, K., et al., 2018. Shifts in plant distributions in response to climate warming in a biodiversity hotspot, the Hengduan Mountains. *J. Biogeogr.* 45, 1334–1344.
- Liang, M., Chen, W., LaFontaine, A.M., et al., 2023. Taxon-specific, phased siRNAs underlie a speciation locus in monkeyflowers. *Science* 379, 576–582.
- Lischer, H.E., Excoffier, L., 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299.
- Liu, J., Möller, M., Provan, J., et al., 2013. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytol.* 199, 1093–1108.
- Liu, B., Abbott, R.J., Lu, Z., et al., 2014. Diploid hybrid origin of *Ostryopsis intermedia* (Betulaceae) in the Qinghai-Tibet Plateau triggered by quaternary climate change. *Mol. Ecol.* 23, 3013–3027.
- Liu, Y.Y., Jin, W.T., Wei, X.X., et al., 2019. Cryptic speciation in the Chinese white pine (*Pinus armandii*): implications for the high species diversity of conifers in the Hengduan Mountains, a global biodiversity hotspot. *Mol. Phylogenet. Evol.* 138, 114–125.
- Liu, J., Milne, R.I., Zhu, G.F., et al., 2022. Name and scale matters: clarifying the geography of Tibetan Plateau and adjacent mountain regions. *Global Planet. Change* 215, 103893.
- Luo, D., Yue, J.P., Sun, W.G., et al., 2016. Evolutionary history of the subnival flora of the Himalaya-Hengduan Mountains: first insights from comparative phylogeography of four perennial herbs. *J. Biogeogr.* 43, 31–43.
- Ma, Y., Mao, X., Wang, J., et al., 2022. Pervasive hybridization during evolutionary radiation of *Rhododendron* subgenus *Hymenanthes* in mountains of southwest China. *Natl. Sci. Rev.* 9, nwa276.
- Mai, D., 2000. Die mittelmiozänen und obermiozänen Floren aus der Mueroer und Raunoer Folge in der Lausitz. Teil I: farnpflanzen, Koniferen und Monokotyledonen. *Palaeontogr. Abt. B Palaeophytol.* 256, 1–68.
- Malinsky, M., Svardal, H., Tyers, A.M., et al., 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecol. Evol.* 457, 830.
- Malinsky, M., Matschner, M., Svardal, H., 2021. Dsuite-fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21, 584–595.
- Mallet, J., 2007. Hybrid speciation. *Nature* 446, 279–283.
- Marchese, C., 2015. Biodiversity hotspots: a shortcut for a more complicated concept. *Glob. Conserv.* 3, 297–309.
- Muchhala, N., Johnsen, S., Smith, S.D., 2014. Competition for hummingbird pollination shapes flower color variation in Andean Solanaceae. *Evolution* 68, 2275–2286.

- Muellner-Riehl, A.N., 2019. Mountains as evolutionary arenas: patterns, emerging approaches, paradigm shifts, and their implications for plant phylogeographic research in the Tibeto-Himalayan Region. *Front. Plant Sci.* 10, 195.
- Muellner-Riehl, A.N., Schnitzler, J., Kissling, W.D., et al., 2019. Origins of global mountain plant biodiversity: testing the 'mountain-geobiodiversity hypothesis'. *J. Biogeogr.* 46, 2826–2838.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., et al., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., et al., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
- Nocchi, G., Wang, J., Yang, L., et al., 2023. Genomic signals of local adaptation and hybridization in Asian white birch. *Mol. Ecol.* 32, 595–612.
- Pertierra, L.R., Segovia, N.I., Noll, D., et al., 2020. Cryptic speciation in gentoo penguins is driven by geographic isolation and regional marine conditions: unforeseen vulnerabilities to global change. *Divers. Distrib.* 26, 958–975.
- Phillips, S.J., Anderson, R.P., Dudík, M., et al., 2017. Opening the black box: an open-source release of Maxent. *Ecography* 40, 887–893.
- Purcell, S., Neale, B., Todd-Brown, K., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.
- Qu, X.J., Moore, M.J., Li, D.Z., et al., 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* 15, 50.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Raj, A., Stephens, M., Pritchard, J.K., 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589.
- Rieseberg, L.H., Willis, J.H., 2007. Plant speciation. *Science* 317, 910–914.
- Rosenberg, N.E., 2004. DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138.
- Schenk, J.J., Axel, J., 2016. Consequences of secondary calibrations on divergence time estimates. *PLoS One* 11, e0148228.
- Schley, R.J., Twyford, A.D., Pennington, R.T., 2022. Hybridization: a 'double-edged sword' for Neotropical plant diversity. *Bot. J. Linn. Soc.* 199, 331–356.
- Schumer, M., Rosenthal, G.G., Andolfatto, P., 2014. How common is homoploid hybrid speciation? *Evolution* 68, 1553–1560.
- Seregin, A.P., Anačkov, G., Friesen, N., 2015. Molecular and morphological revision of the *Allium saxatile* group (Amaryllidaceae): geographical isolation as the driving force of underestimated speciation. *Bot. J. Linn. Soc.* 178, 67–101.
- Struck, T.H., Feder, J.L., Bendiksby, M., et al., 2018. Finding evolutionary processes hidden in cryptic species. *Trends Ecol. Evol.* 33, 153–163.
- Sun, H., Zhang, J., Deng, T., et al., 2017. Origins and evolution of plant diversity in the Hengduan Mountains, China. *Plant Divers.* 39, 161–166.
- Sun, S.S., Guo, Y.L., Favre, A., et al., 2022. Genetic differentiation and evolutionary history of two medicinal gentians (*Gentiana stipitata* Edgew. and *Gentiana szechenyii* Kanitz) in the Qinghai-Tibet Plateau. *J. Appl. Res. Med. Aroma.* 30, 100375.
- Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577.
- Tarasov, A., Vilella, A.J., Cuppen, E., et al., 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 31, 2032–2034.
- Vallejo-Marín, M., Hiscock, S.J., 2016. Hybridization and hybrid speciation under global change. *New Phytol.* 211, 1170–1187.
- van Wyk, A.M., Dalton, D.L., Hoban, S., et al., 2017. Quantitative evaluation of hybridization and the impact on biodiversity conservation. *Ecol. Evol.* 7, 320–330.
- Wan, D.S., Feng, J.J., Jiang, D.C., et al., 2016. The Quaternary evolutionary history, potential distribution dynamics, and conservation implications for a Qinghai-Tibet Plateau endemic herbaceous perennial, *Anisodus tanguticus* (Solanaceae). *Ecol. Evol.* 6, 1977–1995.
- Wang, T.R., Meng, H.H., Wang, N., et al., 2023. Adaptive divergence and genetic vulnerability of relict species under climate change: a case study of *Pterocarya macroptera*. *Ann. Bot.* 132, mcad083.
- Warnock, R.C.M., Parham, J.F., Joyce, W.G., et al., 2015. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proc. Roy. Soc. B-Biol. Sci.* 282, 20141013.
- Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- White, O.W., Reyes-Betancort, J.A., Chapman, M.A., et al., 2020. Geographical isolation, habitat shifts and hybridisation in the diversification of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *New Phytol.* 228, 1953–1971.
- Wright, S., 1943. Isolation by distance. *Genetics* 28, 114–138.
- Wu, S., Wang, Y., Wang, Z., et al., 2022. Species divergence with gene flow and hybrid speciation on the Qinghai-Tibet Plateau. *New Phytol.* 234, 392–404.
- Xing, Y., Ree, R.H., 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci. U.S.A.* 114, E3444–E3451.
- Yang, F.S., Qin, A.L., Li, Y.F., 2012. Great genetic differentiation among populations of *Meconopsis integrifolia* and its implication for plant speciation in the Qinghai-Tibetan Plateau. *PLoS One* 7, e37196.
- Yang, R., Folk, R., Zhang, N., et al., 2019. Homoploid hybridization of plants in the Hengduan mountains region. *Ecol. Evol.* 9, 8399–8410.
- Yuan, Y.M., Küpfer, P., 1997. The monophyly and rapid evolution of *Gentiana* sect. *Chondrophyllae* Bunge s.l. (Gentianaceae): evidence from the nucleotide sequences of the internal transcribed spacers of nuclear ribosomal DNA. *Bot. J. Linn. Soc.* 123, 25–43.
- Zhang, X.L., Wang, Y.J., Ge, X.J., et al., 2009. Molecular phylogeny and biogeography of *Gentiana* sect. *Cruciata* (Gentianaceae) based on four plastid DNA datasets. *Taxon* 58, 862–870.
- Zhang, X., Sun, Y., Landis, J.B., et al., 2020. Genomic insights into adaptation to heterogeneous environments for the ancient relictual *Circaea agrestis* (Circaceasteraceae, Ranunculales). *New Phytol.* 228, 285–301.
- Zhang, Z., Daly, J.S., Tyrrell, S., et al., 2021. Formation of the three gorges (Yangtze River) no earlier than 10 Ma. *Earth Sci. Rev.* 216, 103601.
- Zheng, B.X., Xu, Q.Q., Shen, Y.P., 2002. The relationship between climate change and Quaternary glacial cycles on the Qinghai-Tibetan Plateau: review and speculation. *Quat. Int.* 97–98, 93–101.
- Zheng, W., Yan, L.J., Burgess, K.S., et al., 2021. Natural hybridization among three *Rhododendron* species (Ericaceae) revealed by morphological and genomic evidence. *BMC Plant Biol.* 21, 1–12.