

Conservation Genetics and Modeling Potential Geographic Distribution of *Corybas taliensis*, a Small 'Sky Island' Orchid Species in China

Yuhang Liu

University of Chinese Academy of Sciences, Beijing 100049, China

Huichun Wang

University of Chinese Academy of Sciences, Beijing 100049, China

Jing Yang

Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences (CAS), Kunming, Yunnan 650201, China

Zhiling Dao

Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences (CAS), Kunming, Yunnan 650201, China

Weibang Sun (✉ wbsun@mail.kib.ac.cn)

Kunming Botanical Garden, Kunming Institute of Botany, Chinese Academy of Sciences (CAS), Kunming, Yunnan 650201, China

Research Article

Keywords: *Corybas taliensis*, ddRAD-seq, population genetic, MaxEnt, suitable habitats, conservation management

Posted Date: June 30th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3067811/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Plant Biology on January 2nd, 2024. See the published version at <https://doi.org/10.1186/s12870-023-04693-y>.

Abstract

Background

Corybas taliensis is an endemic species of sky islands in the Pan-Tibetan Plateau region of China. Its habitat is fragile and volatile, and it is likely that the species is threatened. However, it is difficult to determine the conservation priority or unit without knowing the genetic background and the distribution of this species. In this study, we used double digest restriction-site associated DNA-sequencing (ddRAD-seq) to investigate the conservation genomics of *C. taliensis*. At the same time, we modeled the present and future (2030 and 2090) suitable habitats for *C. taliensis* using the maximum-entropy (MaxEnt) model.

Results

The results suggested that the related *C. fanjingshanensis* belongs to *C. taliensis* and should not be considered a separate species. All the sampling locations were divided into three genetic groups: Sichuan & Guizhou population (SG population), Hengduan Mountains population (HD population) and Himalaya population (HM population), and there was complex gene flow between the sampling locations of HD population. MT was distinct genetically from other sampling locations due to the special environment in Motuo. The genetic diversity (π , H_d) of *C. taliensis* was relatively high, but its contemporary effective population size (N_e) was small. *C. taliensis* might be currently affected by inbreeding depression, while its large population density may be able to reduce the effect. The areas of suitable habitats in higher mountains are not predicted to change significantly in the future, and these suitable habitats is predicted to spread to other higher mountains under future climate change. However, suitable habitats in relatively low altitude mountains may disappear in the future. Which indicated that *C. taliensis* will be caught in a 'summit trap' in low altitude areas. On the contrary, the high altitude of Himalaya and Hengduan Mountains regions play as 'biological refuges' for *C. taliensis*.

Conclusions

These results not only provide a new understanding of the genetic background and potential resource distribution of *C. taliensis*, but also lay the foundation for its conservation and management.

Background

Habitats deterioration and destruction caused by anthropogenic stresses and global climate change threaten global ecological function [1, 2, 3], and result in the continued decline of the world 's biodiversity [4, 5]. A thorough assessment of the risks facing biodiversity is needed to plan action to slow rates of species and biodiversity decline and protect and manage biological resources [6]. Endangered species may be more likely to become extinct under global climate change and anthropogenic disturbance, and therefore require more attention and research in a timely manner to prevent their loss [7, 8]. In addition to reproductive difficulties [9], small populations of endangered species are also affected by genetic drift, which can erode the genetic diversity of populations and can lead to a rapid increase in genetic differentiation among populations [10, 11]. The loss of genetic diversity may reduce the persistence of populations, as well as reducing the evolutionary potential of

the species as a whole, and high genetic differentiation may lead to the decline of outbreeding depression [12, 13, 14]. The ecological resilience (the ability to resist or recover from disturbance, also called ecological robustness) and the ability of organisms to adapt to environmental changes are also affected by the underlying genetic characteristics of their populations [15]. Therefore, conservation genetics research is also important for the development of conservation priorities and management strategies for endangered species. Meanwhile, understanding the geographical distribution of threatened taxa is a crucial step towards their conservation [16], and information regarding the potential geographical distributions of species based on biological-climate relationships can help with decision-making in plant protection projects [17].

Concepts of island biogeography, as originally proposed by MacArthur and Wilson (1967), and their later development gave rise to the idea of “sky islands” to describe continental habitats (e.g., mountains) that are isolated from similar habitats by unfavorable terrain (e.g., valleys) [18], in the same way that the sea separates oceanic islands [19]. Species in sky islands face unsuitable climate in the intervening valleys which restricts movement from one montane area to another [20]. This means that the sky islands are not well connected and that there is little migration of individuals through the habitat barrier [21], preventing gene flow between populations in different sky islands, and eventually meaning that these populations have high levels of inter-population genetic divergence and unique patterns of genetic structure [19, 22]. Flightless species in different sky islands are particularly structured [23, 19], as it is very difficult for them to migrate between these mountain habitats [24], although occasional long-distance dispersal events do sometimes occur, possibly mediated by wind or birds [25]. Because species responses to climatic changes are influenced by interacting factors such as ecology, landscape topography, and latitude and longitude, the pattern and tempo of diversification will vary [26]. For example, where a species is present on multiple sky islands, this could result from the fragmentation of a wide-ranging common ancestor or may be due to inter-island dispersal [27]. Within the sky islands themselves, changing climatic conditions can cause suitable environments for species or populations to shift, expand or contract along elevational gradients [28, 29]. This allows these islands to serve as refugia and/or to generate strong ecological gradients [30], and that in turn may allow long-term population persistence during global climate fluctuations in the mountains [31]. Montane habitats therefore often become hotspots of endemism, intraspecific genetic diversity and species richness, and provide opportunities to study how different evolutionary processes lead to diversification [32]. Sky island systems are thus attractive study areas for investigating how heterogeneous landscapes promote population genetic differentiation and speciation, both spatially and temporally [33, 34].

Corybas Salisb. is a genus in the Orchidaceae, originally published by the British botanist Richard Anthony Salisbury in 1807. The type species is *C. aconitiflorus*, from Australia, New Zealand and New Guinea. The scientific name of the genus was chosen because of the resemblance of the corolla of *Corybas* species to the helmet of the mythical Corybas, a priest of the Greek goddess Cybele [35]. The *Flora Reipublicae Popularis Sinicae* (FRPS) classified the genus *Corybas* to the Orchidaceae, Subfam. Orchidoideae, Trib. Neottieae, Subtrib. Acianthinae. A molecular systematics study published in 2015 reclassified Subtrib. Acianthinae as Trib. Diurideae [36].

All species of *Corybas* are small and have bulbous underground tubers, and most of them have unique veins on their leaves and lips. The plants of this genus tend to inhabit relatively conservative microclimates, generally growing in moss or litter on damp and shaded soil slopes or stone slopes [37]. The peculiar growth form and

flowers of *Corybas* give them great ornamental value and development potential as cultivated plants [38, 39]. About 160 species of *Corybas* are known globally, and are mainly distributed from the Himalaya in the north to Macquarie Island, which is about half way between Australia and Antarctica (GBIF, <https://www.gbif.org/>). *Corybas* has centers of diversity in New Guinea, Australia and New Zealand, and China is the northern margin of the genus [40]. Studies on taxonomy [37, 41], population genetics [42] and phylogeny [43, 44, 45] of *Corybas* have analyzed the phylogenetic relationships between different species, and also evaluated the taxonomic delimitation and conservation status in some species. In 2020, [46]. studied the chloroplast genome of *C. taliensis* and found that the species studied with the closest relationship with *Corybas* was *Rhizanthella gardneri*. Through the pollination studies of different species of *Corybas* by Han et al (2022) and Wang (2021), the observed pollinators of *Corybas* have been shown to be small flies of the family Mycetophilidae [47, 48]. Study of the roots of *C. macranthus* from the Catlins region of South Island, New Zealand has demonstrated that the roots contain fungal vesicles of the type found in vesicular-arbuscular (phycomycetous) mycorrhizae [49].

C. taliensis was first published in 1951 by Tang and Wang. The type specimens were collected from Dali, Yunnan, and are now deposited in the Museum of Biological Specimens, Peking University, Beijing (PEY; PEY0051469; Tang and Wang, 1951). Through field investigations, we found nine wild distribution locations of *C. taliensis* in Tibet, Yunnan, Sichuan and Guizhou, which were narrowly restricted to isolated montane regions, mainly between elevations of 2300 and 3500 m. *C. taliensis* can therefore be considered as a sky island species [50]. Most plants were found in clumps of moss under coppiced montane moss forests, or coniferous and broad-leaved mixed forests, with a few plants found growing in the deciduous layer with thick humus or wet rotten wood (Fig. 1). The relative fragility of such habitats, coupled with their location on isolated and complex sky islands, makes the species more vulnerable to stochastic effects and local extinctions [51]. *C. taliensis* has been categorized an Endangered (EN) in both the *Threatened Species List of China's Higher Plants* [52] and the *Chinese Red List of biodiversity-vascular plant volume* (2020), and has been listed as second-class protected plant of the state in *the National key protected wild plants list* published by China in 2021 [53].

Corybas taliensis in the wild. The habitat (A), habitat and flowers (B, C), and flower (D).

In this study, ddRAD-seq was selected to conduct conservation genetics research on *C. taliensis*. ddRAD-seq is widely used in plant population genetics, phylogenetics, phylogeography and molecular breeding [54, 55] as a simple, fast and inexpensive method of constructing RAD-Seq databases [56]. We used the MaxEnt [57] to predict the current and future distribution of habitats that is potentially suitable for *C. taliensis*, because MaxEnt can deal with presence-only data and has been shown to outperform most other models, even the ensemble model [58, 59], and it performs better with small sample sizes relative to other species distribution models [60, 61]. The objectives of this study are (1) to explore the population structure in *C. taliensis*, to infer the levels of gene flow among populations, and to estimate the genetic diversity and N_e of each population; (2) to analyze the possible reasons underlying the observed genetic patterns in *C. taliensis* as a sky island species; (3) to use the MaxEnt to predict areas of potentially suitable habitats for *C. taliensis* under current and future climate predictions. We then aim to combine the results of our genetic analysis with the prediction of suitable habitats, and to use these data to suggest conservation and management suggestions.

Results

Processing of Sequencing Data

A total of 542 million reads were generated from all samples in 9 sampling locations (Dingjie, Zhari, Motuo, Pianma, Tengchong, Eryuan, Cangshan, Puge, Guizhou sampling locations; abbreviated as: DJ, ZR, MT, PM, TC, EY, CS, PG, GZ), and after filtering out low-quality reads and reads without RAD-tags, 528 million reads remained for processing. The sequencing depth of samples ranged from 5.90× (CS05) to 28.80× (GZ14), with an average coverage of 13.32×. We used Stacks to carry out de novo analysis on *C. taliensis* [62], through SNP calling with the values of the parameters M, m and n, set to M = 1, m = 2, n = 2, which resulted in 5116 single nucleotide polymorphisms (SNPs; Table S1). Tajima's D statistical test showed a negative average Tajima's D values (-0.369) with a high degree of significance (p-value < 0.05) for 5116 SNPs, which 4222 were significantly negative (Fig. S1).

Population Structure and Genetic Diversity

Principal component analysis (PCA) analysis supported the separation of all samples into three groups (Fig. 2), ZR and DJ clustered together and were defined as HM population, while PG and GZ clustered together and were defined as SG population. All the other sampling locations (CS, EY, TC, MT and PM) made up HD population. The relationship between HM population and HD population was close, and both fell relatively far away SG population in the PCA analysis. From the CV error values, the optimal *K* values included 2, 3 and 4 (Fig. S2). When *K* = 2, DJ, ZR and MT formed one genetic group, and the other sampling locations (PM, TC, EY, CS, PG and GZ) formed another one. When *K* = 3, DJ and ZR formed one genetic group, and the other two groups were formed by MT, PM, TC, EY and by CS, PG, GZ. When *K* = 4, CV error values was the lowest, DJ, ZR and PG, GZ formed two independent genetic groups, while PM, TC, EY and CS formed a genetic group (named "YN population", as all samples were from Yunnan Province), and with MT on its own as a group (Fig. 3). This result was similar to the results from the PCA, but with the main difference that in the case of Bayesian cluster analysis (*K* = 4), MT was suggested to be an independently segregated group with a small amount gene flow between YN population (PM, TC, EY, CS), while in the PCA, MT grouped together with CS, EY, TC and PM. BayesAss analysis showed that there was no obvious recent gene flow among sampling locations (Table 1). Mantel tests of geographic distance and genetic differentiation showed that isolation-by-distance (IBD) was an important factor for the genetic differentiation of *C. taliensis* ($R^2 = 0.372$, $P < 1e-2$) (Fig. S3). The fixation index (F_{ST}) between DJ and PG was 0.121, and that between DJ and GZ was 0.114 (Table 2). AMOVA analysis showed that 0.67 % of genetic variation could be found in individuals from SG population, HM population and HD population from PCA analysis; -0.23% of the genetic variation between the four genetic groups (SG, HM, MT, YN populations) from Bayesian clustering analysis; -2.09% genetic variation was found among the sampling locations (Table S2).

Table 1 Recent migration rates between each sampling location

Sampling location	PG	DJ	ZR	PM	TC	EY	GZ	MT	CS
PG	0.8727	0.0159	0.0159	0.0160	0.0160	0.0161	0.0158	0.0156	0.0161
DJ	0.0168	0.8663	0.0168	0.0165	0.0168	0.0167	0.0166	0.0166	0.0169
ZR	0.0152	0.0154	0.8780	0.0154	0.0152	0.0151	0.0151	0.0153	0.0152
PM	0.0154	0.0158	0.0157	0.8736	0.0160	0.0164	0.0156	0.0159	0.0156
TC	0.0147	0.0147	0.0145	0.0145	0.8838	0.0143	0.0141	0.0147	0.0146
EY	0.0211	0.0208	0.0207	0.0210	0.0205	0.8335	0.0207	0.0208	0.0207
GZ	0.0146	0.0145	0.0147	0.0145	0.0144	0.0143	0.8842	0.0144	0.0146
MT	0.0139	0.0137	0.0140	0.0139	0.0140	0.0139	0.0138	0.8888	0.0139
CS	0.0137	0.0140	0.0140	0.0141	0.0138	0.0139	0.0136	0.0139	0.8892

Table 2 Genetic differentiation coefficient (F_{ST}) between each sampling location

Sampling location	DJ	ZR	PM	TC	EY	GZ	MT	CS
PG	0.121	0.099	0.112	0.102	0.109	0.084	0.110	0.082
DJ		0.080	0.094	0.085	0.104	0.114	0.090	0.090
ZR			0.078	0.072	0.083	0.094	0.073	0.080
PM				0.056	0.080	0.107	0.076	0.074
TC					0.069	0.098	0.069	0.064
EY						0.104	0.088	0.075
GZ							0.103	0.079
MT								0.078

The number of individuals (N) investigated in the wild in DJ and TC was the most (N = 9000), EY (N = 5000) and CS (N = 6000) had moderate counts, and there were relatively fewer at PM, MT, ZR, PG and GZ (N <1000) (Table 3). In the population genetic statistics, the nucleotide diversity (π) values were not significantly different among groups, and ranged from 0.0723 (CS) to 0.1031 (ZR) at the level of sampling locations, from 0.0939 (SG population) to 0.1025 (HM population) at populations level, and 0.1011 at species level. The observed heterozygosity (H_o) ranged from 0.0308 (PM) to 0.0553 (CS) for the different sampling locations, from 0.0355 (HM population) to 0.0471 (SG population) at the level of genetic group and 0.0418 at species level. The expected heterozygosity (H_e) values at location, genetic group and species levels were 0.0690 (CS)–0.0972 (ZR), 0.0915 (SG population)–0.0996 (HM population) and 0.1005, respectively. The inbreeding coefficients (F_{IS}) were greater than 0 at all levels, and ranged from 0.0677 (CS) to 0.2142 (ZR), from 0.1985 (HM population) to 0.4052

(HD population) and from 0.5878 at the three levels, respectively. NeEstimator V2.1 was not able to obtain reliable N_e estimates (inf) for ZR, PM, TC, EY, GZ, MT and CS, but the N_e estimates for the other four sampling locations were 16.3 (PG), 17.7 (DJ), 76.4 (MT) and 14.5 (CS). The estimated N_e values in the three genetic groups were 4.7 (SG population), 64.8 (HM population) and 23.4 (HD population). The N_e estimate for species was 28.7 (Table 3).

Table 3 Summary of genetic diversity, effective population size and the number of individuals

Genetic groups	Sampling locations	Samples	H_o	H_e	π	F_{IS}	$N_e(95\%C.L.)$	N
SG	-	30	0.0471	0.0915	0.0939	0.1985	4.7 (2.8, 7.7)	1260
	PG	15	0.0467	0.0757	0.0799	0.1047	16.3 (2.5, inf)	560
	GZ	15	0.0397	0.0867	0.0919	0.1745	Inf (15.5, inf)	700
HM	-	30	0.0355	0.0996	0.1025	0.3222	64.8 (16.3, inf)	9500
	DJ	15	0.0495	0.0738	0.0780	0.0918	17.7 (2.7, inf)	9000
	ZR	15	0.0396	0.0972	0.1031	0.2142	Inf (9.9, inf)	500
HD	-	75	0.0421	0.0954	0.0964	0.4052	23.4 (16.9, 33.0)	20414
	PM	15	0.0308	0.0789	0.0839	0.1712	Inf (31.0, inf)	107
	TC	15	0.0359	0.0825	0.0870	0.1723	Inf (17.1, inf)	9000
	EY	15	0.0341	0.0766	0.0834	0.1435	Inf (1.6, inf)	5000
	MT	15	0.0421	0.0784	0.0830	0.1291	76.4 (10.8, inf)	307
	CS	15	0.0553	0.0690	0.0723	0.0677	14.5 (4.9, 235.4)	6000
all	-	135	0.0418	0.1005	0.1011	0.5878	28.7 (23.5, 34.9)	31174

Data Preparation for Distribution Prediction

After processing and screening, we found that 28 distribution points were suitable in the prediction of suitability of habitats (Fig. 4). Ten environmental factors were selected for MaxEnt modeling, including four climate factors: (isothermality (bio3), temperature seasonality (bio4), temperature annual range (bio7), precipitation of driest month (bio14)), and six soil factors (sand percentage of topsoil (T_SAND), silt percentage of topsoil

(T_SILT), bulk density of topsoil (T_BULK_DEN), organic carbon percentage of topsoil (T_OC), total exchangeable base of topsoil (T_TEB), bulk density of subsoil (S_BULK_DEN)).

Prediction of the Potential Geographical Distribution

The area under the curve (AUC) training value was 0.991 (Fig. S4), and all known distribution points in predicted suitable habitat range, indicating that the model performed well and generated reliable evaluations. According to the current suitable habitat distribution map, the habitats with only low suitability for *C. taliensis* was mainly predicted to be found in the Himalaya, the Hengduan Mountains, Yunnan, Taiwan and Hainan, with a few areas in the cities along the southern coast of China. Moderately suitable habitats were predicted mainly in Yunnan and Taiwan, with a few areas in southern Guangdong and the southern Himalaya in China. The highly suitable habitats were predicted mainly in Taiwan and Hainan (Fig. 5A). The predicted distribution of habitats with low suitability in the Himalaya and the Hengduan Mountains will spread continually northwards towards the Tibetan Plateau in 2030 and 2090 under the ssp126 climate scenario, and the low suitable habitats in the south coastal area will increase first and then decrease. Moderately suitable habitats were predicted to decrease in the north of Yunnan in 2030 and increase in the Himalayan region in 2090 (Fig. 5B, C). Under the ssp585 climate scenario, the change in the predicted distribution of suitable habitats in 2030 compared with current estimates, only the low suitable habitats in the Himalaya and the Hengduan Mountains will spread to the north, other areas had no obvious change. In 2090, habitats with low suitability will spread further northwards into the Tibetan Plateau than in 2030, low suitable habitats in the central almost disappear, and the moderately suitable habitats will be more reduced in Yunnan than in 2030, while moderately suitable areas in the Himalaya and Hengduan Mountains will spread northward (Fig. 5D, E). There was never any significant change in the predicted distribution of highly suitable habitats.

In terms of area change (Table S3 and Fig. S5A-C), under the ssp126 climate scenario, the area of habitats with low suitability and that of highly suitable habitats are predicted to increase in 2030. By 2090, the area of highly suitable habitats is predicted to shrink to below the current level, and the area of habitats of low suitability will shrink but still be higher than the current level. The area of moderately suitable habitats is predicted to have decreased by 2030 but then to increase again between 2030 and 2090, but to still be lower than the current level. Under the ssp585 climate scenario, the area of habitats with low suitability is predicted to have increased by 2030 and again by 2090, while the areas of moderately and highly suitable habitats are predicted to initially decrease and then to increase slightly, but these areas are not predicted to recover to the present extent.

Discussion

Population Structure

The earliest described *Corybas* from Guizhou (GZ) was described as a new species, *C. fangjingshanensis* [40]. In this study, the inference of genetic structure from the ddRAD-seq demonstrated that *C. fangjingshanensis* was not an independent species, but instead should be synonymized with *C. taliensis* (Figs. 2, 3).

The results of our AMOVA analysis supported the grouping from the PCA analysis (Table S2), and for this reason, the analysis of population structure was mainly based on the results of the PCA analysis, with the

Bayesian clustering analysis ($K = 4$) used as auxiliary information. Mantel tests of geographic and genetic distances between our nine sampling locations indicated a significant pattern of IBD in *C. taliensis* (Fig. S3). There was little genetic differentiation among the sampling locations (Table 2), with the most obvious differentiation being between DJ and GZ. The geographic distance between them was the farthest of all the sampling locations (Fig. 6), which may be one of the causes of this differentiation. The negative Tajima's D values of *C. taliensis* might indicate the genes with an excess of low-frequency variation, which was consistent with the fact that the populations might have experienced an expansion after a recent bottleneck, or the genes targeted are under positive selective pressure [63].

Map of all sampling locations. Except for PG and GZ, all of the other sampling locations are in the Pan-Tibetan Highlands. ZR, DJ are in Himalaya, and CS, EY, TC, MT and PM are all from the Hengduan Mountains.

Except for SG population, which is located in the high mountains of Sichuan (PG: Ganhaizi 2300–3200 m a.s.l.) and Guizhou (GZ: Fanjing Mountain 2300–2320 m a.s.l.), all of the other sampling locations are in the Pan-Tibetan Highlands [64, 65]. HM population (ZR, DJ) is in Himalaya, and HD population (CS, EY, TC, MT and PM) is from the Hengduan Mountains (Fig. 6). The Himalaya-Hengduan Mountains (HHM) region is a significant reservoir of global biodiversity in general [16] and for alpine plant life in particular [66]. Unlike the continuous range of the Himalaya, the Hengduan Mountains constitute a series of high, parallel, mostly north–south running ridges, separated by deep valleys [67] forming groups of sky islands. There was no obvious recent gene flow among the sampling locations, which might suggest that the genetic patterns observed in *C. taliensis* populations are more likely to be as a result of ancestral distribution rather than to gene flow: e.g., *Drosophila innubila* [24]. The hypothesis that niche conservatism drives diversification in sky island species predicts that lineages should be separated by unsuitable conditions [68, 21, 69], and sky island populations are therefore predicted to have high levels of inter-population genetic divergence and unique patterns of genetic structure [19, 68].

DJ and ZR of HM population are two distant sampling locations on the same sky island. According to the idea that niche conservatism drives diversification in sky island species, the observation that these two populations clustered together in a single group may result from the similar environmental characteristics of the two sampling locations on the same mountain, leading to similar levels of genetic differentiation [70]. GZ and PG of SG population are far away from each other in spatial distance, however, they obviously clustered into one group in the PCA analysis (Fig. 2), which contradicts both the result of the Mantel test (Fig. S3) and the hypothesis of diversification in sky island species. PG and GZ in SG population are located on different sky islands, with PG located in the south of Sichuan and GZ located in the northeast of Guizhou. However, the Yangtze River flows through the two locations [71], Heishui River, a first-class tributary of Jinshajiang River (part of the upstream of Yangtze River) flows through PG, and Yinjiang River, a tributary of Wujiang River (a first-class tributary of Yangtze River), flows from east to west from the south side of GZ. Recent genetic analyses suggests that intercatchment (and upstream) dispersal does occur in many plant species [72, 73, 74, 75], so we speculate that the Yangtze River flowing through these two localities may somehow connect them, or perhaps it is that the environmental characteristics of the two localities is similar enough that it leads to similar levels of genetic differentiation [70]. The sampling locations of HD population (CS, EY, TC, MT and PM) were all located on different mountains of the Hengduan sky islands group, and the gene flow between was complex (Fig. 3). This may be because the north–south mountain chains and valleys can provide corridors for floral exchange

between the north and south but represent barriers to migration between the east and west [70, 76, 77]. Most sky island species are highly divergent between locations from different sky islands [23, 78, 79], and our population structure results therefore suggests that the *C. taliensis* populations from the Hengduan sky islands group may be much more recently established than other studied species. The current lack of geographic divergence may be due to the maintenance of ancestral variation. If this is the case, we expect that the *C. taliensis* sampling locations will become as structured as those of other species over time [24].

HD population lies midway between SG population and HM population, and there were small amounts of geneflow diffusion on both sides towards SG population and HM population. It may be that these populations share a common ancestor, or perhaps because there is occasionally dispersal over long distances mediated by wind or birds. MT lies in Motuo and is relatively distant from the other sampling locations of HD population. It is quite distinct genetically, although the Bayesian cluster analysis ($K = 4$) suggested that this population had a small amount gene flow with YN population (PM, TC, EY, CS). Motuo crosses many different climate and vegetation zones, and the climate of this area is complex with large vertical variation [80]. The unique geography and climate type have allowed this area to evolve a unique orchid flora [81], which may explain why MT is genetically distinct from all the other sampling locations. MT is located in the western edge of the Hengduan Mountains near the Himalaya. The PCA analysis suggested an aggregation of MT and YN population (Fig. 2). This may be due to gene flow carried by the channels formed by the north-south position of the Hengduan Mountains, while the same mountains form a barrier between MT and HM population, blocking their east-west communication.

Genetic Diversity

Estimation of N_e from genetic marker data is a major focus for biodiversity conservation because it allows estimation of the rate at which inbreeding is increasing and at which additive genetic variation is lost [82]. The N_e values of ZR, PM, TC, EY and GZ sampling locations were infinite (Table 3), which may have resulted from either a truly large N_e or as a consequence of limited data quality [83]. Low N_e values and low heterozygosity indicate that species have low environmental suitability, that their long-term survival was challenging and that they are at high risk of extinction [84, 85, 86]. The low N_e of *C. taliensis* indicated that its ability to adapt to changing environments might be low. Some studies have suggested that the minimum N_e needed to avoid inbreeding depression is 100 [87], but the valid levels of N_e in this study did not reach 100 in any group (Table 3), meaning that *C. taliensis* might be currently affected by inbreeding depression. In the study of ddRAD sequencing analyses in several orchid species [88], the diversity levels of orchids with high genetic diversity described in Roy et al. (2017) were lower than that of *C. taliensis* (Table 3) in our study, so we suggest that the genetic diversity of *C. taliensis* is therefore relatively high, which conflicted with the low N_e . Populations may be able to maintain genetic variation by density-dependent selection [89], and, for this reason, may be less vulnerable to the detrimental effects of inbreeding [90]. The N from the sampling locations in our field investigation lying between 107 and 9000 (Table 3), which may could explain the conflict between high genetic diversity and low N_e of *C. taliensis*: the large population density may be able to reduce the effect of inbreeding depression.

Predicting the Potential Geographical Distribution

The predicted areas of suitable habitats for *C. taliensis* in the present and future are mainly found in Yunnan, the mountains of Taiwan, the Hengduan Mountains and the Himalaya (Fig. 5A-E). The ranges of elevationally determined species have been shown to change with climate change [91]. On mountains around the world, increasing temperatures force alpine plants to migrate upwards until they reach the highest elevations [92]. The higher the mountain, the more future refuges for organisms under future climate change [93, 94]. The Pan-Tibetan Plateau has experienced significant warming in recent decades, which has led to large-scale glacier contraction [95, 96], lake formation and expansion [97] and extensive evolution of vegetation [98]. There are a large number of mountains with an altitude of about 4000–5000 m distributed in the Tibetan Plateau [65], and continuous global warming will lead to further contraction of glaciers. It is likely that this will provide suitable habitats for the migration of *C. taliensis*, and is likely to be the reason that habitats with low suitability for *C. taliensis* in the Himalaya and Hengduan Mountains regions is predicted to spread northward into the Tibetan Plateau in the future (Fig. 5B-D, E).

Unfortunately, with climate change, Species growing in high altitude areas are often caught in ‘summit traps’: because species inhabiting mountain summits are forced to move to higher elevations with temperature increase, with further increases in temperature they may have no escape route and may become locally extinct [92]. The distribution of moderately suitable habitats in Yunnan is likely to contract in the future (Fig. 5E), and it may be that *C. taliensis* will be caught in a ‘summit trap’ here. In contrast, the moderately suitable habitats in the Himalaya and Hengduan Mountains regions are predicted to continue to spread, possibly due to the high altitude of these areas play as ‘biological refuges’. The distribution of highly suitable habitats changed very little (Fig. 5A-E), which may because it becomes increasingly difficult to find highly suitable habitats for the conservative *C. taliensis* under global climate change and human disturbance.

Conclusions

Dispersal in *C. taliensis* is difficult and, as it is a sky island species, the future distribution of its habitats is unstable. However, the species has relatively high genetic diversity. Nine sampling locations were divided into three genetic groups: SG population, HD population and HM population. Sampling locations of *C. taliensis* from different sky island were not highly structured, this might be related to the maintenance of ancestral variation, the similar environmental characteristics or the media transmission of wind or birds. The gene flow of HD population was more complicated than other populations. MT was distinct genetically from other sampling locations due to the special environment in Motuo. With their limited geographic range, endemic species of sky islands are particularly susceptible to climate change [99]. The result of modeling potential geographic distribution indicated that *C. taliensis* will be caught in a ‘summit trap’ in low altitude areas. On the contrary, the high altitude of Himalaya and Hengduan Mountains regions play as ‘biological refuges’ for *C. taliensis*. Management and protection plan for *C. taliensis* need to take into account the different climate predicted for different areas and over different periods in the future. In our field investigations, we found that some sampling locations (including PG and EY) were affected by grazing and that the *C. taliensis* plants were also sometimes collected by local people. Both of these activities can threaten the survival of plant species. At the same time, because of its low environmental adaptability and narrow habitats range, it is not feasible to do ordinary cultivation experiments on *C. taliensis*, and simple ex-situ conservation strategies are therefore currently unrealistic.

The following suggestions are put forward for the protection and management of *C. taliensis*. 1) Mountain peaks behave like islands, and show high levels of genetic isolation, so we suggest that the locations of *C. taliensis* should be managed like islands, for instance promoting the use of native germplasm for regeneration efforts [31]. TC and CS are in nature reserves and are likely to be well protected, but other locations lie outside nature reserves. We suggest that part of each sky island should be designated a protected area for the targeted protection and management of this species, and effective population protection monitoring should be continually carried out to prevent local extinctions caused by human disturbance and environmental fluctuations. 2) The environment and habitats preferred by *C. taliensis* should be investigated in detail, and we should aim to create a suitable environment for cultivation in an area that is convenient and reachable, so that ex-situ cultivation and conservation can be carried out. After this ex-situ conservation technology is realized, germplasm resources from PM, MT, ZR, PG, GZ and EY, all of which are greatly disturbed by humans or have only very small numbers of individuals, can be collected and grown ex situ, to ensure that the resources are preserved before the species becomes extinct at these localities. Our population structure analyses suggest that SG population (GZ, PG), HM population (ZR, DJ) and HD population (CS, EY, TC, MT and PM) represent different genetic backgrounds, and the germplasm resources from these different genetic groups need to be collected and preserved. The individuals at YN population may face local extinction because of the ‘summit trap’, and the protection and preservation of the germplasm resources from this location is a priority. The MT population represents a unique genetic group, and it is necessary to focus on the preservation of the individuals from this location, to preserve the diversity genetic resources of this species. 3) The areas of potentially suitable habitats predicted in our study should be used to look for more potential distribution of *C. taliensis* in the wild, which can then be used to protect and manage further resources. 4) The populations in the Himalaya and Hengduan Mountains have the potential to spread further northward into the Tibetan Plateau in the future. We should investigate whether these populations have indeed spread northwards in these areas through long-term population and environmental monitoring, and should make timely investigation and protection. We believe that if the above protection and management measures are implemented, the likelihood of long-term survival of *C. taliensis* will be increased.

Methods

Sampling, DNA Extraction and Sequencing

A total of 135 individuals of *C. taliensis* from 9 locations (DJ, ZR, MT, PM, TC, EY, CS, PG, GZ) were sampled (Fig. 6), with 15 individuals being sampled from each location. The interval distance between individuals was at least 10 m. The environmental characteristics were recorded and the N at each sampling location was estimated while sampling. Huichun Wang conducted the plant material collection and the formal identification for this study. One voucher specimen was collected for each sampling location and deposited in the Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming (code HW2020001–HW2020009, see Table S1). The collection of leaves and voucher specimens of *C. taliensis* in this study was permitted by the Motuo National Nature Reserve, Cangshan and Erhai National Nature Reserve Management Center and Gaoligongshan National Nature Reserve. The collecting of all materials complied with the Regulations of the People’s Republic of China on the Protection of Wild Plants and the IUCN Policy Statement on Research Involving Species at Risk of Extinction.

Total DNA was extracted from leaf tissue dried in silica gel using the cetyl trimethylammonium bromide (CTAB) method [100]. The quantity and quality of each sample was determined using agarose gel (Omega Bio-Tek, Norcross, Ga., United States) electrophoresis and on a fluorometer (Qubit 3.0, Thermo Fisher Scientific, Waltham, Ma., United States). Each qualified DNA sample was standardized to the same volume (10 μ L) and quantity (200 ng). Then, 10 μ L of pre-mixed double enzyme (EcoRI and MseI; New England Biolabs, Ipswich, Ma., United States) digestion solution was added to each sample. Samples were then placed in a PCR machine (Takara PCR Thermal Cycler Dice, Takara Bio Inc., Shiga, Japan) at 37°C for 8 h, followed by 65°C for 20 min and with a final incubation step at 12°C. The digested fragments were detected using agarose gel electrophoresis, and were then ligated to EcoRI and MseI adapters containing sample specific barcodes using a PCR machine under the following conditions: 16°C for 8 h, 65°C for 20 min, and 12°C for the final incubation. Barcoded samples were separated by size (350–500 bp) using an agarose gel. To meet the concentration requirements for sequencing, each library was then amplified using a PCR machine under the following conditions: 98°C for 2 min, followed by 25 cycles for 98°C for 15 s, 56°C for 15 s and 72°C for 15 s; and 72°C for 5 min. The paired-end sequencing was performed on an Illumina HiSeq X-Ten platform (Illumina Inc., San Diego, CA, United States).

Processing of Next Generation Sequencing Data

Stacks v2.62 [101, 62] was used to process the reads generated by ddRAD-seq. Samples were initially demultiplexed and filtered with `process_radtags`, with `len_limit` set to 140 bp to trim low-quality reads, and `retain_header -t` set to 135. Next, the `ustacks`, `cstacks`, `sstacks`, `tsvbam`, `gstacks`, and `populations` modules were run to call SNPs. We used different combinations of the parameters `M` (Maximum distance in nucleotides allowed between stacks), `m` (Minimum depth of coverage required to create a stack), and `n` (number of mismatches allowed between sample loci when build the catalog) for SNP calling, and finally chose the best combination of parameters by screening through the maximum number of polymorphic sites in many individuals. This process includes (i) selecting the optimal `m` value in the range of 2–7 when default `M` = 2 and `n` = 0, (ii) selecting the optimal `M` value in the range of 1–5 with the `m` value optimized in the previous step and `n` = 0, and (iii) using the previously optimized `M` and `m` values, selecting the optimal `n` value among `M` – 1, `M` and `M` + 1 [102]. The resulting SNP data sets were used for subsequent analysis, during which the required data format conversion was performed using PGDSpider v2.1.1.5 [103]. Tajima's *D* values were calculated by VCFtools v0.1.16 in 500-kb windows to test all the loci for neutrality [104].

Population Structure and Genetic Diversity

PCA and Bayesian clustering were used to deduce the population structure of *C. taliensis*. PCA analysis was performed using the R package "adegenet" [105], and confidence ellipses were drawn for each sample location including 95% of the individuals. The Admixture software was used for the Bayesian clustering analysis [106]. To determine the optimal value of *K*, we tested the number of clusters from 1 to 9 and estimated the optimal number of clusters as that with the lowest cross-validation error rate. In R v4.2.2 [107], the packages "Pophelper" [108] and "Bar Plots" function [109] were used to visualize the Admixture results. Recent gene flow between sampling locations was assessed using the Bayesian inference approach BayesAss [110] as implemented in the BA3SNP program [111], and was performed over 1.0×10^7 iterations, with a burn-in of 1.0×10^6 steps, and sampled at a frequency of 1000 times. AMOVA analysis was carried out in Arlequin v3.5.2.2 [112] to calculate

genetic differences among species, genetic groups, and sampling locations. To test for IBD, pairwise geographical distances among the 9 sampling locations were computed using GENALEX v.6.5 [113], with pair wise genetic distance taken to be $F_{ST}/(1-F_{ST})$. Mantel tests were performed in R using the package "vegan" [114], with 10,000 permutations.

The π , H_e , H_o , F_{ST} and F_{IS} were calculated for each sampling location, genetic group and species using the populations module in the Stacks pipeline. The linkage disequilibrium method in NeEstimator V2.1 [83] was used to estimate the N_e for each sample location and for each population (sample size = n), and a cutoff was set within the following interval: $1/(2n) \leq \text{PCRIT} \leq 1/n$ [115].

Data Preparation for Distribution Prediction

The occurrence records of *C. taliensis* were obtained through the Chinese Virtual Herbarium (CVH; <https://www.cvh.ac.cn/index.php>), the Flora of China (www.iplant.cn/foc/), official notices, field investigation data and consultation with experts. Duplicate records were deleted. If a record had only fuzzy, county-level information, a site was chosen using the Baidu Coordinate Picking System (<https://api.map.baidu.com/lbsapi/getpoint/>) according to the known information, and was then optimized in LocaSpace Viewer v4 (<http://www.locaspace.cn/LSV.jsp>) according to the known topographic characteristics of the *C. taliensis* habitats. The distribution data were screened in ArcGIS 10.8 to ensure that the distances between the final distribution points were all more than 1 km.

Initially, we selected 55 climatic factors and 44 soil factors that may have an impact on the present and future distribution of habitats suitable for *C. taliensis*. All climate data were downloaded from WorldClim v2.1 (www.worldclim.org): 19 bioclimatic variables, 12 months of monthly total precipitation, 12 months of monthly average maximum temperature and 12 months of monthly average minimum temperature. We downloaded the relevant data for the years 1970–2000 as current climate data, and chose future climate data as predicted by scenarios that modeled either sustainable development with lower emissions (SSP126) or conventional development with higher emissions (SSP585) from the CNRM-ESM2-1 global climate change model [116]. Soil data was downloaded from Harmonized World Soil Database v1.2 [117]. All data sets had a resolution of 30 s. In ArcGIS 10.8, all environmental data were clipped to China, and the number of rows and columns, cell size, extent and spatial reference of all layers were checked for consistency. In order to avoid multicollinearity of variables that can result in model over-fitting [118], we used Pearson correlation analysis to examine the cross-correlation between environmental factors, and removed one of each pair of environmental factors where the Pearson correlation coefficient was greater than 0.75 and p-value was less than 0.05.

Predicting the Potential Geographical Distribution

We used the MaxEnt v3.4.4 [119] to simulate potential habitats suitable for *C. taliensis* using the following parameters: random test percentage: 25%, regularization multiplier: 1, replicated run type: bootstrap, and with 10 repeats. The final prediction result gives a suitability range of 0–1 for the potentially suitable habitats, and we reclassified these in ArcGIS 10.8 as follows: 0–0.01 was defined as unsuitable habitats, 0.01–0.35 was habitats with low suitability, 0.35–0.65 was defined as moderately suitable habitats and 0.65–1 was defined as highly

suitable habitats. The area of habitats at each level of suitability was calculated under the projected coordinate system 'WGS_1984_UTM_Zone_47N', and a line chart of area change was drawn.

Abbreviations

ddRAD-seq

double digest restriction-site associated DNA-sequencing

MaxEnt

maximum-entropy model

SG population

Sichuan & Guizhou population

HD population

Hengduan Mountains population

HM population

Himalaya population

YN population

Yunnan population

DJ

Dingjie sampling location

TC

Tengchong sampling location

EY

Eryuan sampling location

ZR

Zhari sampling location

MT

Motuo sampling location

PG

Puge sampling location

GZ

Guizhou sampling location

CS

Cangshan sampling location

PM

Pianma sampling location

IBD

isolation-by-distance

N

number of individuals

F_{ST}

fixation index

π

nucleotide diversity

H_o

observed heterozygosity

H_e

expected heterozygosity

F_{IS}

inbreeding coefficients

N_e

contemporary effective population size

Declarations

Ethics approval and consent to participate:

The authors confirm that the plant material from *C. taliensis* used are available. The experimental research on plants complied with relevant institutional, national, and international guidelines and legislation. The collection of leaves and voucher specimens of *C. taliensis* in this study was permitted by the Motuo National Nature Reserve, Cangshan and Erhai National Nature Reserve Management Center and Gaoligongshan National Nature Reserve. The collecting of all materials complied with the Regulations of the People's Republic of China on the Protection of Wild Plants and the IUCN Policy Statement on Research Involving Species at Risk of Extinction, and the collection of all plant materials does not pose any risk to other species in nature.

Consent for publication:

Not applicable.

Availability of data and materials:

The datasets analysed during the conservation genetics study are available in the National Center for Biotechnology Information (NCBI) repository, <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA972527>. The other datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests:

The authors declare that they have no competing interests.

Funding

This work was supported by the Second Tibetan Plateau Scientific Expedition and Research Program [2019QZKK0502], and by the Science and Technology Basic Resources Investigation Program of China "Survey and Germplasm Conservation of Plant Species with Extremely Small Populations in Southwest China" [2017FY100100].

Authors' contributions:

WS conceived and designed the work. YL performed the statistical analyses, interpreted the results, and wrote the manuscript. HW collected the leaf materials. HW, JY and ZD revised the manuscript. WS acquired the

funding. All authors read and approved the final manuscript.

Acknowledgements:

We would like to thank Lidan Tao, Yulong Yu, Pin Zhang, and Lei Cai for help with sampling, and Fengmao Yang for his help with data analysis.

Authors' information:

Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences (CAS), Kunming, Yunnan 650201, China

Yuhang Liu, Huichun Wang, Jing Yang, Zhiling Dao & Weibang Sun

Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences (CAS), Kunming, Yunnan 650201, China

Jing Yang, Zhiling Dao & Weibang Sun

Kunming Botanical Garden, Kunming Institute of Botany, Chinese Academy of Sciences (CAS), Kunming, Yunnan 650201, China

Weibang Sun

University of Chinese Academy of Sciences, Beijing 100049, China

Yuhang Liu & Huichun Wang

References

1. Hoffmann M, Hilton-Taylor C, Angulo A, et al. The impact of conservation on the status of the world's vertebrates. *Science*. 2010;330:1503.
2. Venter O, Sanderson EW, Magrath A, et al. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat Commun*. 2016;7:12558.
3. Powers RP, and Jetz W. Global habitats loss and extinction risk of terrestrial vertebrates under future land-use change scenarios. *Nat Clim Change*. 2019;9:323–329. doi:10.1038/s41558-019-0406-z
4. Vitousek PM, Mooney HA, Lubchenco J, et al. Human domination of Earth's ecosystems. *Science*. 1997;277:494–499.
5. Butchart SHM, Walpole M, Collen B, et al. Global biodiversity: indicators of recent declines. *Science*. 2010;328:1164–1168.
6. Ayensu E, Claasen DR, Collins M, et al. International ecosystem assessment. *Science*. 1999;286:685–686.
7. Bijlsma R, and Loeschcke V. Genetic erosion impedes adaptive responses to stressful environments. *Evol Appl*. 2012;5:117–129. doi:10.1111/j.1752-4571.2011.00214.x
8. Ralls K, Ballou JD, Dudash MR, et al. Call for a paradigm shift in the genetic management of fragmented populations. *Conserv Lett*. 2018;11:12412. doi:10.1111/conl.12412

9. Kéry M, Matthies D, and Spillmann HH. Reduced fecundity and offspring performance in small populations of the declining grassland plants *Primula veris* and *Gentiana lutea*. J Ecol. 2000;88:17–30. doi:10.1046/j.1365-2745. 2000.00422.x
10. Willi Y, Van BJ, Schmid B, et al. Genetic isolation of fragmented populations is exacerbated by drift and selection. J Evol Biol. 2007;20:534–542. doi:10.1111/j.1420-9101. 2006.01263.x
11. Haag T, Santos AS, Sana DA, et al. The effect of habitats fragmentation on the genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (*Panthera onca*). Mol Ecol. 2010;19:4906–4921. doi:10.1111/j.1365-294X. 2010. 04856.x
12. Elam E. Population genetic consequences of small population size: implications for plant conservation. Annu Rev Ecol Syst. 1993;24:217–242. doi:10.1146/annurev.es.24.110193.001245
13. Spielman D, Brook BW, and Frankham R. Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci. U.S.A. 2004;101:15261–15264. doi:10.1073/pnas.0403809101
14. Bomblies K, and Weigel D. *Arabidopsis*: a model genus for speciation. Curr Opin Genet Dev. 2007;17:500–504. doi:10.1016/j.gde.2007.09.006
15. Wernberg T, Coleman MA, Bennett S, et al. Genetic diversity and kelp forest vulnerability to climatic stress. Sci Rep. 2018;8:1851.
16. Myers N, Mittermeier RA, Mittermeier CG, et al. Biodiversity hotspots for conservation priorities. Nature. 2000;403:853–858. doi:10.1038/35002501
17. Yan X, Wang S, Duan Y, et al. Current and future distribution of the deciduous shrub *Hydrangea macrophylla* in China estimated by MaxEnt. Ecol Evol. 2021;11:16099–16112. doi:10.1002/ece3.8288
18. Gorman ML. Continental habitats islands. In: Gorman ML, editor. Island ecology. Outline studies in ecology, Springer. Dordrecht, Netherlands. 1979.p.53–64.
19. McCormack JE, Huang H, and Knowles LL. Sky islands. In: McCormack JE, Huang H, and Knowles LL, editors. Encyclopedia of islands. Berkeley, CA: University of California Press, Berkeley. 2009;4,pp.841–843.
20. Sekar S and Karanth P. Flying between Sky Islands: The Effect of Naturally Fragmented Habitats on Butterfly Population Structure. PLoS One. 2013;8:e71573. doi:10.1371/journal.pone.0071573
21. Wiens JJ and Graham CH. Niche conservatism: integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst. 2005;36:519–539.
22. Wu YK, Wang YZ, Jiang K, et al. Significance of pre-quaternary climate change for montane species diversity: insights from Asian salamanders (Salamandridae: *Pachytriton*). Mol Phylogenet Evol. 2013;66:380–90.
23. Favé MJ, Johnson RA, Cover S, et al. Past climate change on sky islands drives novelty in a core developmental gene network and its phenotype. BMC Evol Biol. 2015;15:183.
24. Hill T and Unckless RL. Adaptation, ancestral variation and gene flow in a 'sky island' *Drosophila* species. Mol Ecol. 2021;30:83–99. doi:10.1111/mec.15701
25. Assefa A, Ehrich D, Taberlet P, et al. Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. Heredity. 2007;99:133–142. doi:10.1038/sj.hdy.6800974
26. Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc B. 2004;359:183–195.

27. Shepard DB and Burbrink FT. Lineage diversification and historical demography of a sky island salamander, *Plethodon ouachitae*, from the Interior Highlands. *Mol Ecol.* 2008;17:5315-5335. doi:10.1111/j.1365-294X.2008.03998.x
28. Shepard DB and Burbrink FT. Local-scale environmental variation generates highly divergent lineages associated with stream drainages in a terrestrial salamander, *Plethodon caddoensis*. *Mol Phylogenet Evol.* 2011;59:399–411.
29. Herman F, Champagnac JD. Plio-Pleistocene increase of erosion rates in mountain belts in response to climate change. *Terra Nova.* 2016;28:2–10.
30. Hedin M, Carlson D and Coyle F. Sky island diversification meets the multispecies coalescent – divergence in the spruce-fir moss spider (*Microhexura montivaga*, Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. *Mol Ecol.* 2015;24: 3467-3484. doi:10.1111/mec.13248
31. Mastretta-Yanes A, Xue AT, Moreno-Letelier A, et al. Long-term in situ persistence of biodiversity in tropical sky islands revealed by landscape genomics. *Mol Ecol.* 2018;27:432– 448. doi:10.1111/mec.14461
32. Taubmann J, Theissinger K, Feldheim KA, et al. Modelling range shifts and assessing genetic diversity distribution of the montane aquatic mayfly *Ameletus inopinatus* in Europe under climate change scenarios. *Conserv Genet.* 2011;12:503–15.
33. Knowles LL and Richards CL. Importance of genetic drift during Pleistocene divergence as revealed by analyses of genomic variation. *Mol Ecol.* 2005;14:4023–4032.
34. Cox SC, Prys-Jones RP, Habel JC, et al. Niche divergence promotes rapid diversification of East African sky island white-eyes (Aves: Zosteropidae). *Mol Ecol.* 2014;23:4103–4118.
35. Arditti J and Van RP. *Corybas*. *Bull Torrey Bot Club.* 1984;111:231–232.
36. Chase MW, Cameron KM, Freudenstein JV, et al. An updated classification of Orchidaceae. *Bot J Linn Soc.* 2015;177:151–174.
37. Faria E. Diversity in the genus *Corybas* Salisb. (Orchidaceae, Diurideae) in New Caledonia. *Adansonia.* 2016;38:175–198.
38. Tang T and Wang FT. *Corybas* Salisb., a new addition to the orchid flora of China. *Acta Phytotaxon Sin.* 1951;1:185–187.
39. Luo Y. Conservation Strategy of Orchidaceae in China. *Forestry of China.* 2003;22:25 –26.
40. Xiong YX, Luo YC, Shangguan FZ, et al. *Corybas fanjingshanensis* Y.X Xiong, a new species of Orchidaceae from Guizhou, China. *Acta Phytotax Sin.* 2007; 45: 808-812.
41. Go R, Ching TM, Nuruddin AA, et al. Extinction risks and conservation status of *Corybas* (Orchidaceae; Orchidoideae; Diurideae) in Peninsular Malaysia. *Phytotaxa.* 2015;233:273–280.
42. Van EML, Lehnebach CA, Pearson SM, et al. Microsatellite markers for *Corybas* (Orchidaceae) species in New Zealand. *Appl Plant Sci.* 2018;6:e1192.
43. Lehnebach CA, Zeller AJ, Frericks J, et al. Five new species of *Corybas* (Diurideae, Orchidaceae) endemic to New Zealand and phylogeny of the Nematoceras clade. *Phylotaxa.* 2016;270:1,1-24. doi:10.11646/phytotaxa.270.1.1
44. Wagner ND, Clements MA, Simpson L, et al. Conservation in the face of hybridisation: genome-wide study to evaluate taxonomic delimitation and conservation status of a threatened orchid species. *Conserv Genet.* 2021;22:151–168. doi:10.1007/s10592-020-01325-y

45. Lyon SP. Molecular Systematics, Biogeography, and Mycorrhizal Associations in the Acianthinae (Orchidaceae), with a Focus on the Genus *Corybas*. The University of Wisconsin, Madison. 2014;161 p.
46. Wang HC, Yang J and Sun WB. Complete chloroplast genome of the endangered *Corybas taliensis* (Orchidaceae), a plant species with extremely small populations endemic to China. Mitochondrial DNA Part B-Resources. 2020;5:1884–1885.
47. Han ZD, Wu Y, Bernhardt P, et al. Observations on the pollination and breeding systems of two *Corybas* species (Diurideae; Orchidaceae) by fungus gnats (Mycetophilidae) in southwestern Yunnan, China. BMC Plant Biol. 2022;22:426 doi:10.1186/s12870-022-03816-1
48. Wang HC. Conservation biology of *Corybas taliensis*. Tang & F. T. Wang. Tutor: Weibang Sun. Kunming Institute of Botany, Chinese Academy of Sciences. 2021.
49. Hall IR. Vesicular mycorrhizas in the Orchid *Corybas macranthus*. Trans Br Mycol Soc. 1976;66:160. doi:10.1016/S0007-1536(76)80108-8
50. Zhang MY, Fritsch PW, Ma PF, et al. Plastid phylogenomics and adaptive evolution of *Gaultheria* series Trichophyllae (Ericaceae), a clade from sky islands of the Himalaya-Hengduan Mountains. Mol Phylogenet Evol. 2017;110:7-18. doi:10.1016/j.ympev.2017.01.015
51. Salerno PE, Señaris JC, Rojas-Runjaic FJM, et al. Recent evolutionary history of Lost World endemics: Population genetics, species delimitation, and phylogeography of sky-island treefrogs. Mol Phylogenet Evol. 2015;82:314-323. doi:10.1016/j.ympev.2014.10.020
52. Qin H, Yang Y, Dong S, et al. Threatened species list of China's higher plants. Biodivers Sci. 2017;25:696–744.
53. Ministry of Ecology and Environment of the People's Republic of China. The Red List of Biodiversity in China-Higher Plants Volume (2020). 2023. <https://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/202305/W020230522536559098623.pdf>. Accessed 19 May 2023.
54. Ulaszewski B, Meger J and Burczyk J. Comparative Analysis of SNP Discovery and Genotyping in *Fagus sylvatica* L. and *Quercus robur* L. Using RADseq, GBS, and ddRAD Methods. Forests. 2021;12:222.
55. Vargas OM, Ortiz EM and Simpson BB. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). New Phytol. 2017;214:1736-1750. doi:10.1111/nph.14530
56. Robledo D, Palaikostas C, Bargelloni L, et al. Applications of genotyping by sequencing in aquaculture breeding and genetics. Rev Aquacult. 2018;10:670-682.
57. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190:231–259. doi:10.1016/j.ecolmodel.2005.03.026
58. Hao T, Elith J, Lahoz-Monfort JJ, et al. Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. Ecography. 2020;43:549–558. doi:10.1111/ecog.04890
59. Kaky E, Nolan V, Alatawi A, et al. A comparison between Ensemble and MaxEnt species distribution modelling approaches for conservation: A case study with Egyptian medicinal plants. Ecol Inform. 2020;60:101150. doi:10.1016/j.ecoinf.2020.101150

60. Pearson RG, Raxworthy CJ, Nakamura M, et al. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J Biogeogr.* 2007;34:102–117. doi:10.1111/j.1365-2699.2006.01594.x
61. Phillips SJ and Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography.* 2008;31:161–175. doi:10.1111/j.0906-7590.2008.5203.x
62. Rochette NC, and Catchen JM. Deriving genotypes from RAD-seq short-read data using Stacks. *Nat Protoc.* 2017;12:2640–2659. doi:10.1038/nprot.2017.123
63. Nedoluzhko A, Orlova SY, Kurnosov DS, et al. Genomic Signatures of Freshwater Adaptation in Pacific Herring (*Clupea pallasii*). *Genes.* 2022;13:1856. doi: 10.3390/genes13101856
64. Liu J and Zhu G. Geographical and geological GIS boundaries of the Tibetan Plateau and adjacent mountain regions (Version 2022.1). 2022. doi:10.5281/zenodo.6432940
65. Liu J, Milne RI, Zhu G, et al. Name and scale matter: Clarifying the geography of Tibetan Plateau and adjacent mountain regions. *Global and Planetary Change.* 2022;215:0921-8181. doi:10.1016/j.gloplacha.2022.103893
66. Wu CY and Wu SG. A proposal for new floristic kingdom (realm)–the E. Asiatic Kingdom, its delineation and characteristics. In: A.L. Zhang and S.G. Wu, editors. *Floristic characteristics and diversity of eastern Asian plants.* Springer-Verlag, Hong Kong. 1996.pp.3–42.
67. Shi YF, Li JJ and Li BY. Uplift and environmental changes of Qinghai-Tibetan Plateau in the Late Cenozoic. Guangdong Science and Technology Press. Guangzhou, China. 1998.
68. Wiens JJ. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution.* 2004;28:193–197.
69. Kozak KH and Wiens JJ. Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution.* 2006;60:2604–2621.
70. Chen YS, Deng T, Zhou Z, et al. Is the East Asian flora ancient or not? *Natl Sci Rev.* 2018;5:920–932.
71. Yangtze River Delta Science Data Center. National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China. 2015. doi:10.12041/geodata.269789738293082.ver1.db
72. Pollux BJA, Jong MDE, Steegh A, et al. Reproductive strategy, clonal structure and genetic diversity in populations of the aquatic macrophyte *Sparganium emersum* in river systems. *Mol Ecol.* 2007;16:313–325.
73. Fér T and Hroudová Z. Detecting dispersal of *Nuphar lutea* in river corridors using microsatellite markers. *Freshwater Biol.* 2008;53:1409–1422.
74. Smulders MJM, Cottrell JE, Lefèvre F, et al. Structure of the genetic diversity in black poplar (*Populus nigra* L.) populations across European river systems: consequences for conservation and restoration. *For Ecol Manage.* 2008;255:1388–1399.
75. Honnay O, Jacquemyn H, Van LK, et al. Temporal and spatial genetic variation in a metapopulation of the annual *Erysimum cheiranthoides* on stony river banks. *J Ecol.* 2009;97:131–141. doi:10.1146/annurev.ecolsys.110308.120159
76. Gao LM, Möller M, Zhang XM, et al. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Mol Ecol.* 2007;16:4684–4698.

77. Li Y, Zhai SN, Qiu YX, et al. Glacial survival east and west of the 'Mekong–Salween Divide' in the Himalaya-Hengduan Mountains region as revealed by AFLPs and cpDNA sequence variation in *Sinopodophyllum hexandrum* (Berberidaceae). *Mol Phylogenet Evol.* 2011;59:412–424.
78. Smith CI and Farrell BD. Phylogeography of the longhorn cactus beetle *Moneilema appressum* LeConte (Coleoptera: Cerambycidae): was the differentiation of the Madrean sky islands driven by Pleistocene climate changes? *Mol Ecol.* 2005;14:3049–3065.
79. Wiens JJ, Camacho A, Goldberg A, et al. Climate change, extinction, and sky island biogeography in a montane lizard. *Mol Ecol.* 2019;28:2610–2624.
80. Wang XL, Li JW, Wang CW, et al. New information of orchids in Tibet, China. *Guihaia.* 2018;38:1440-1445.
81. Lang KY. A study on the geographical distribution and floristic features of the Xizang (Tibet) Orchid flora. *J Univ Chin Acad Sci.* 1980;18:391-407.
82. Ryman N, Laikre L and Hössjer O. Do estimates of contemporary effective population size tell us what we want to know? *Mol Ecol.* 2019;28:1904–1918. doi:10.1111/mec.15027
83. Do C, Waples RS, Peel D, et al. Ne estimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour.* 2014;14:209–214. doi:10.1111/1755-0998.12157
84. Kramer AT and Havens K. Plant conservation genetics in a changing world. *Trends Plant Sci.* 2009;14:599–607.
85. Oakley CG and Winn AA. Effects of population size and isolation on heterosis, mean fitness, and inbreeding depression in a perennial plant. *New Phytol.* 2012;196: 261–270.
86. Poudel RC, Möller M, Liu J, et al. Low genetic diversity and high inbreeding of the endangered yews in Central Himalaya: implications for conservation of their highly fragmented populations. *Diversity and Distributions.* 2014;20:1270–1284.
87. Richard F, Corey JA, Bradshaw BWB. Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation.* 2014;170:56-63.
88. Roy SC, Moitra K and De Sarker D. Assessment of genetic diversity among four orchids based on ddRAD sequencing data for conservation purposes. *Physiol Mol Biol Plants.* 2017;23:169–183. doi:10.1007/s12298-016-0401-z
89. Wallace B. Fifty years of genetic load: An odyssey. Cornell Univ. Press, Ithaca, NY. 1991.
90. Dara N, Diana P. Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. 1997;354–362. doi:10.1111/j.1558-5646.1997.tb02422.x
91. Grabherr G, Gottfried M, and Pauli H. Climate effects on mountain plants. *Nature.* 1994;369:448.
92. Pertoldi C and Bach LA. Evolutionary aspects of climate-induced changes and the need for multidisciplinary. *J Therm Biol.* 2007;32:118–124.
93. Loarie SR, Duffy PB, Hamilton H, et al. The velocity of climate change. *Nature.* 2009; 462:1052–5. doi:10.1038/nature08649
94. Sandel B, Arge L, Dalsgaard B, et al. The influence of Late Quaternary climate-change velocity on species endemism. *Science.* 2011;334:660–664. doi:10.1126/science.1210173

95. Kraaijenbrink PDA, Stigter EE, Yao T, et al. Climate change decisive for Asia's snow meltwater supply. *Nat Clim Chang*. 2021;11:591–597. doi:10.1038/s41558-021-01074-x.
96. Bhattacharya A, Bolch T, Mukherjee K, et al. High Mountain Asian glacier response to climate revealed by multi-temporal satellite observations since the 1960s. *Nat Commun*. 2021;12:4133. doi:10.1038/s41467-021-24180-y
97. Chen F, Zhang M, Guo H, et al. Annual 30 m dataset for glacial lakes in High Mountain Asia from 2008 to 2017. *Earth Syst Sci Data*. 2021;13:741–766. doi:10.5194/essd-13-741-2021
98. Liu Y, Li Z and Chen Y. Continuous warming shift greening towards browning in the Southeast and Northwest High Mountain Asia. *Sci Rep*. 2021;11:17920. doi:10.1038/s41598-021-97240-4
99. Kupfer JA, Balmat J and Smith JL. Shifts in the potential distribution of sky island plant communities in response to climate change. In: Gottfried GJ, Gebow BS, Eskew LG, Edminster CB, Comps. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO. Connecting Mountain islands and desert seas: biodiversity and management of the Madrean Archipelago II. U.S.; 2005.pp. 485–490.
100. Doyle J and Doyle JL. Isolation of plant DNA from fresh tissue. *Focus*. 1990;12:13–15.
101. Catchen J, Hohenlohe PA, Bassham S, et al. Stacks: an analysis tool set for population genomics. *Mol Ecol*. 2013;22:3124–3140. doi:10.1111/mec.12354
102. Díaz-Arce N and Rodríguez-Ezpeleta N. Selecting RAD-Seq Data Analysis Parameters for Population Genetics: The More the Better? *Front Genet*. 2019;10:533. doi:10.3389/fgene.2019.00533
103. Lischer HEL, and Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*. 2011;28:298–299. doi:10.1093/bioinformatics/btr642
104. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123:585–595. doi:10.1093/genetics/123.3.585
105. Jombart T, and Ahmed I. Adegnet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. 2011;27:3070–3071. doi:10.1093/bioinformatics/btr521
106. Alexander DH, Novembre J, and Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–1664. doi:10.1101/gr.094052.109
107. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. 2018.
108. Francis RM. Pophelper: an R package and web app to analyse and visualize population structure. *Mol Ecol Resour*. 2017;17:27–32. doi:10.1111/1755-0998.12509
109. Murrell P. R Graphics. Chapman & Hall/CRC Press. 2005.
110. Wilson GA, and Rannala B. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*. 2003;163:1177–1191. doi:10.1093/genetics/163.3.1177
111. Musmann SM, Douglas MR, Chafin TK, et al. BA3- SNPs: contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods Ecol Evol*. 2019;10:1808–1813. doi:10.1111/2041-210X.13252
112. Excoffier L and Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under linux and windows. *Mol Ecol Resour*. 2010;10:564–567. doi:10.1111/j.1755-0998.2010.02847.X

113. Peakall R and Smouse PE. GenALEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 2012;28:2537–2539.
114. Dixon P. Vegan: a package of R functions for community ecology. *Journal of Vegetation Science*. 2003;14:927-930. doi:10.1111/j.1654-1103.2003.tb02228.x
115. Waples RS, and Do C. Linkage disequilibrium estimates of contemporary ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol Appl*. 2010;3:244–262. doi:10.1111/j.1752-4571. 2009.00104.x
116. Fick SE and Hijmans RJ. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. 2017;37:4302–4315.
117. Fao Soils Portal. Harmonized World Soil Database v 1.2, Fischer, G., Nachtergaele, F., Prieler, S., van Velthuizen, H. T., Verelst, L., and Wiberg, D. 2008. <https://www.fao.org>.
118. Graham MH. Confronting multicollinearity in ecological multiple regression. *Ecology*. 2003;84:2809–2815. doi:10.1890/02-3114
119. Phillips SJ, Dudík M, and Schapire RE. Maxent software for modeling species niches and distributions (Version 3.4.1). 2020. http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed 15 March 2020.

Figures



Figure 1

Corybas taliensis in the wild. The habitat (A), habitat and flowers (B, C), and flower (D).

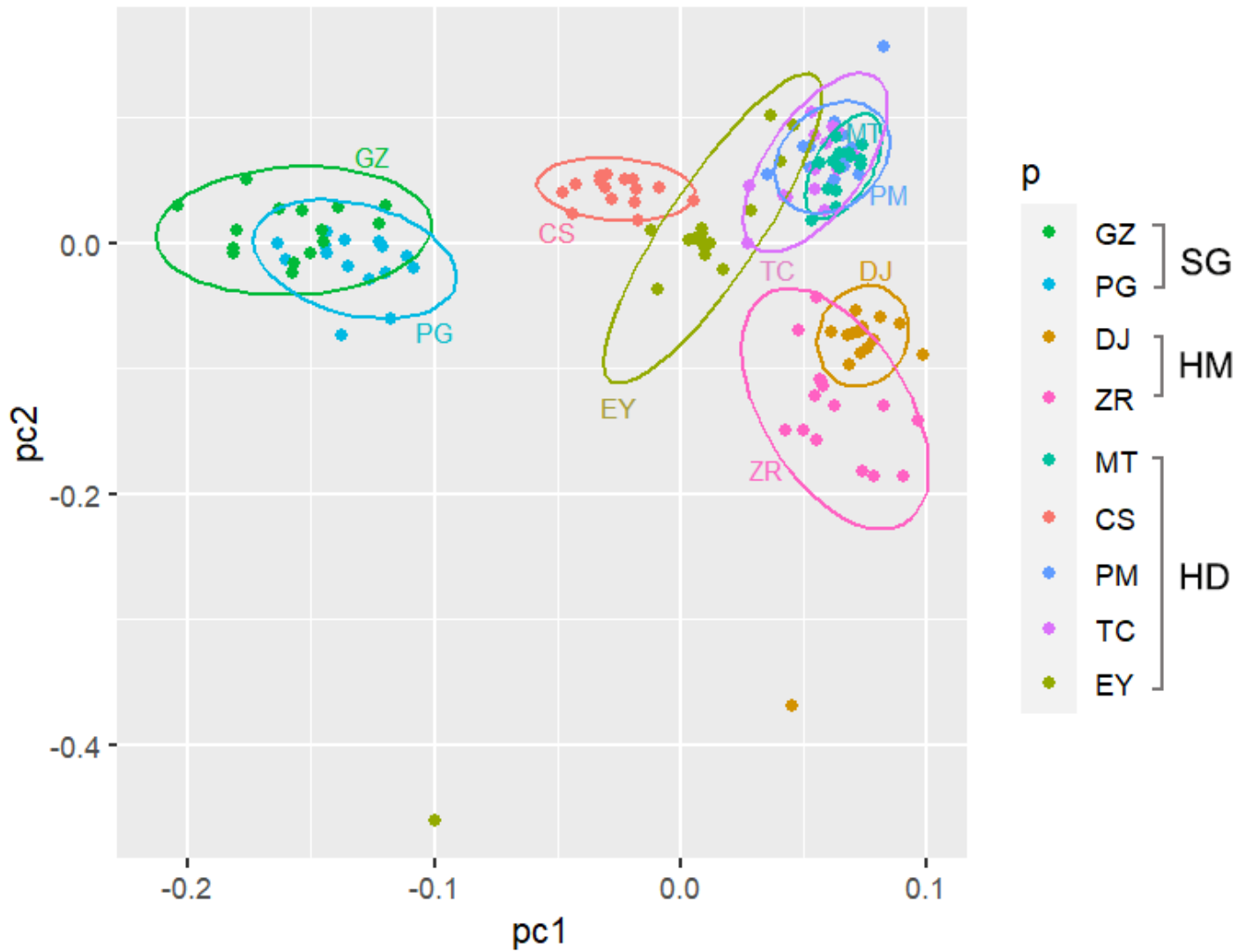


Figure 2

The PCA analysis result. Sampling locations: DJ(Dingjie), ZR(Zhari), MT(Motuo), PM(Pianma), TC(Tengchong), EY(Eryuan), CS(Cangshan), PG(Puge), GZ(Guizhou). Populations: SG (Sichuan & Guizhou population), HD (Hengduan Mountains population), HM (Himalaya population). ZR and DJ clustered together and were defined as HM population, while PG and GZ clustered together and were defined as SG population. All the other sampling locations (CS, EY, TC, MT and PM) made up HD population.

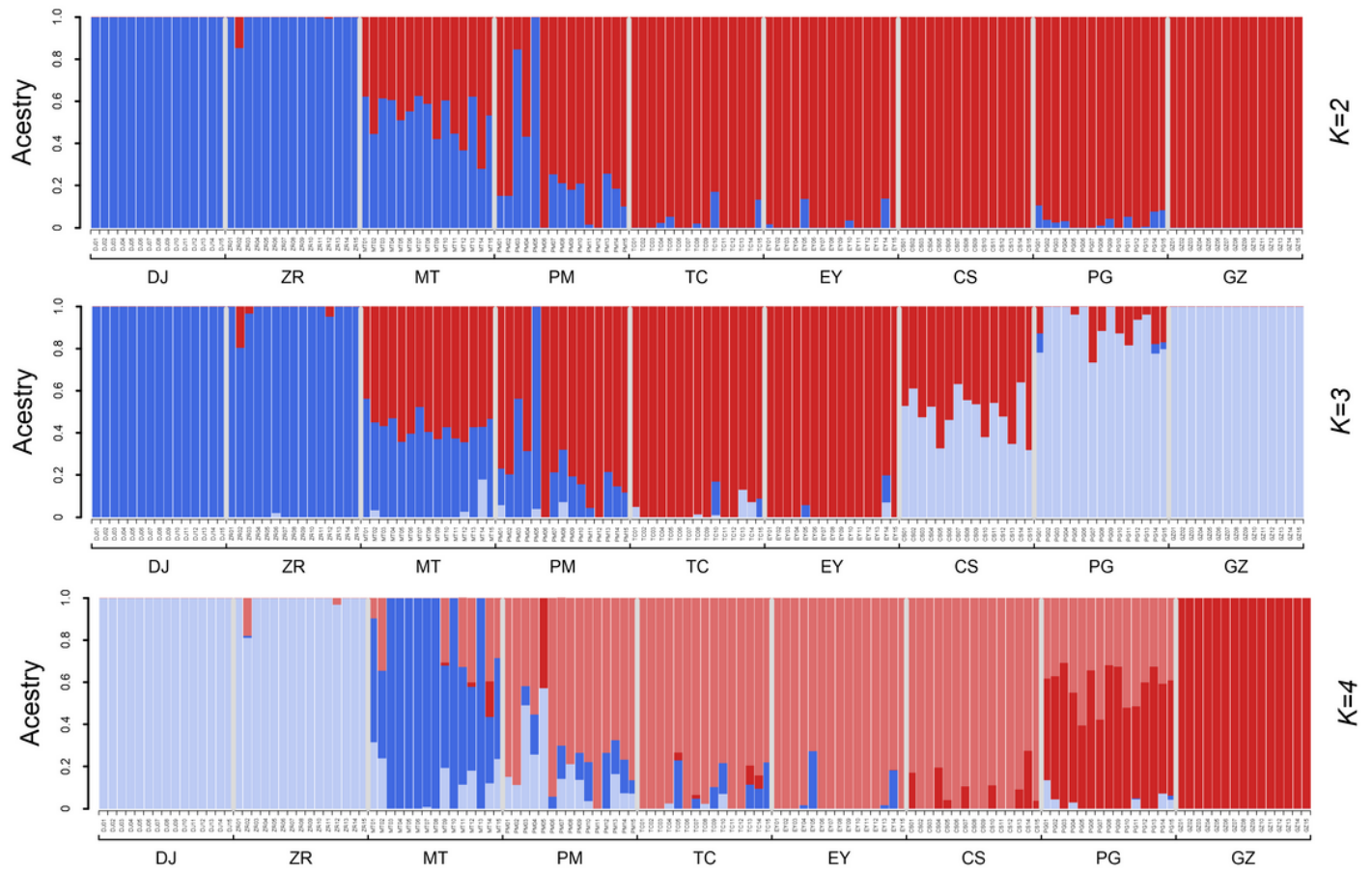


Figure 3

Bayesian cluster analysis result ($K=2$, $K=3$, $K=4$). When $K = 4$, ZR and DJ clustered together and were defined as HM population, while PG and GZ clustered together and were defined as SG population. MT was suggested to be an independently segregated group. All the other sampling locations (CS, EY, TC, MT and PM) made up YN population.

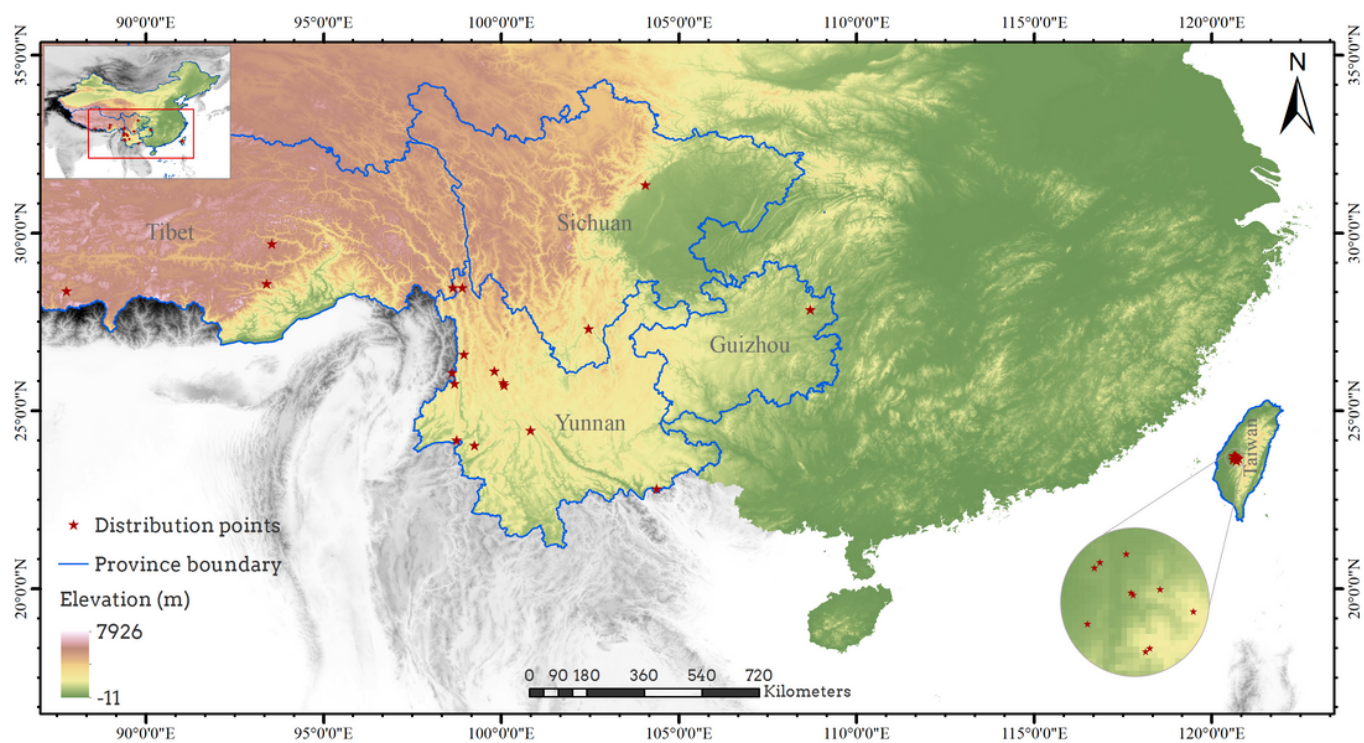


Figure 4

The 28 distribution points of *Corybas taliensis* used for modeling potential geographic distribution.

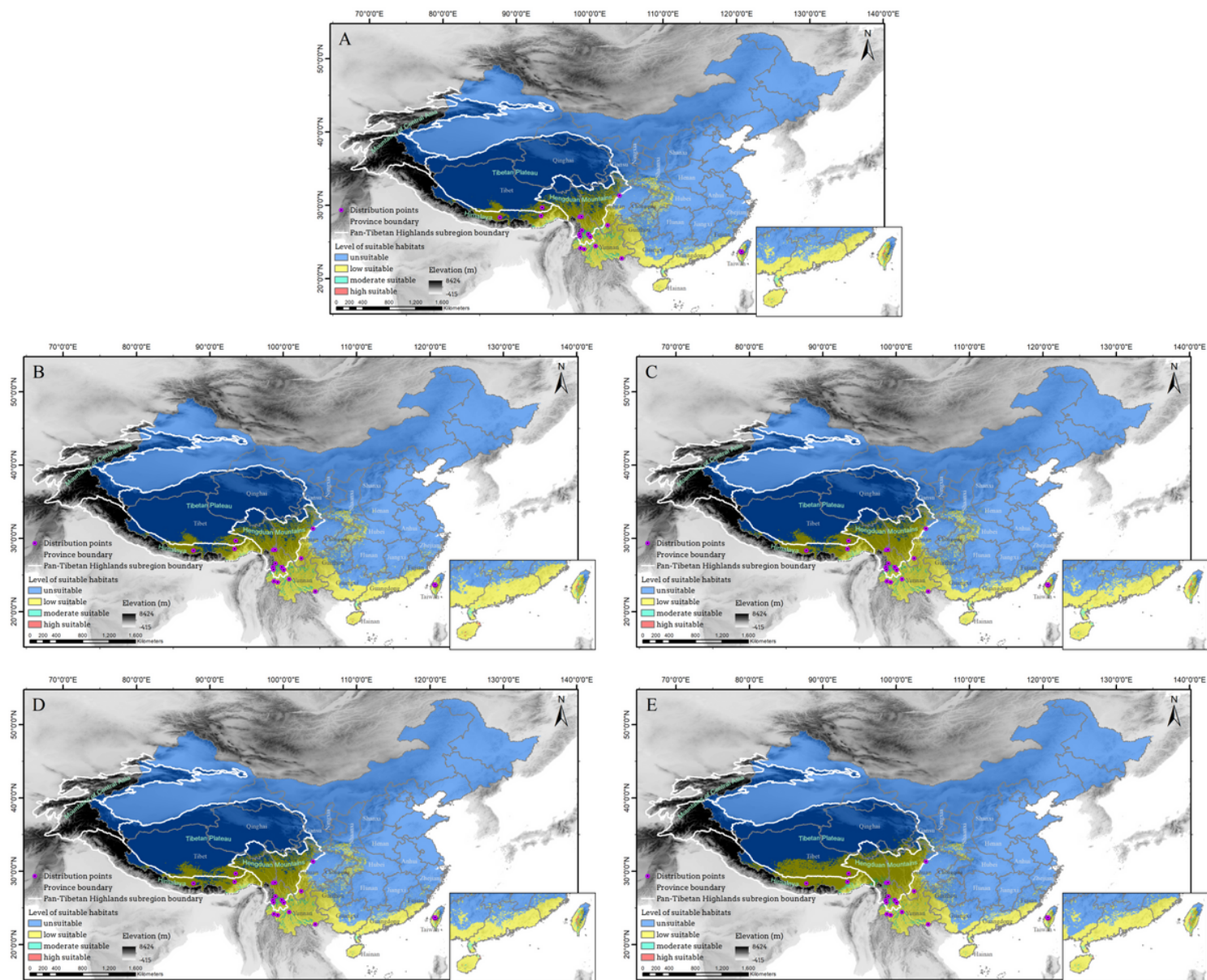


Figure 5

The suitable habitat distribution map. Including the current suitable habitat distribution (A), the suitable habitats distribution in 2030 (B) and 2090 (C) under the ssp126 climate scenario, and in 2030 (D) and 2090 (E) under the ssp585 climate scenario.

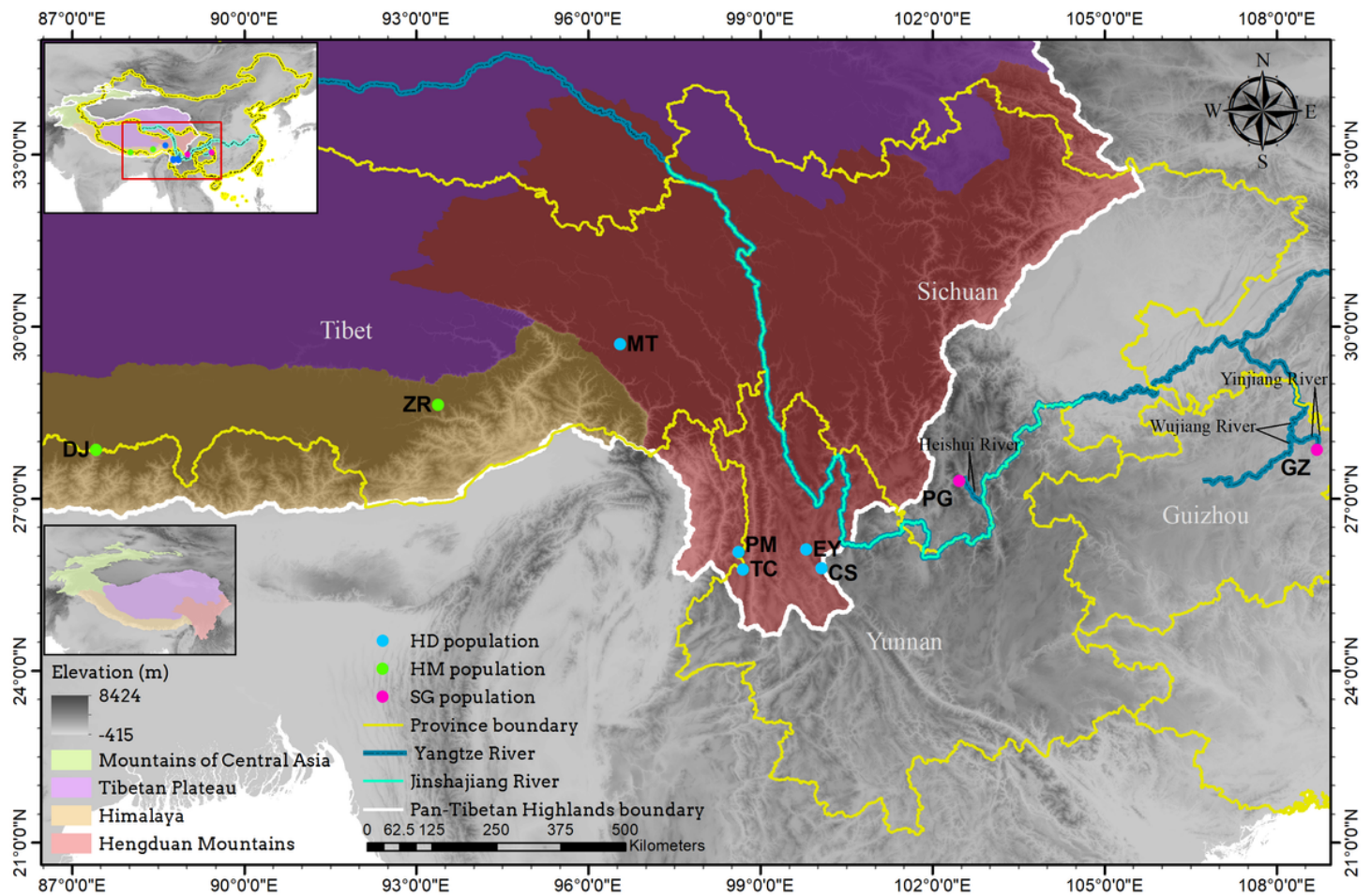


Figure 6

Map of all sampling locations. Except for PG and GZ, all of the other sampling locations are in the Pan-Tibetan Highlands. ZR, DJ are in Himalaya, and CS, EY, TC, MT and PM are all from the Hengduan Mountains.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Fig.S1.png](#)
- [Fig.S2.png](#)
- [Fig.S3.png](#)
- [Fig.S4.png](#)
- [Fig.S5.png](#)
- [Supplementaryfile.docx](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)