Localized environmental heterogeneity drives the population differentiation of two endangered and endemic *Opisthopappus* Shih species

Hang Ye
Shanxi Normal University
Zhi Wang
Shanxi Normal University
Huimin Hou
Shanxi Normal University
Jiahui Wu
Shanxi Normal University
Yue Gao
Shanxi Normal University
Wei Han
Shanxi Normal University
Wenming Ru
Changzhi University
Genlou Sun
Saint Mary's University
Yiling Wang (ylwangbj@hotmail.com)
https://orcid.org/0000-0002-0358-0851

Research article

Keywords: Opisthopappus, genetic differentiation, environmental heterogeneity, Taihang Mountains

Posted Date: February 6th, 2020

DOI: https://doi.org/10.21203/rs.2.22820/v1

Abstract

Background: Climate heterogeneity not only indirectly shapes the genetic structures of plant populations, but also drives adaptive divergence by impacting demographic dynamics. The variable localized climate and topographic complexity of the Taihang Mountains make them a major natural boundary in Northern China that influences the divergence of organisms distributed in this region. *Opisthopappus* is an endemic genus of the Taihang Mountains that includes only two spatially partitioned species *Opisthopappus longilobus* and *Opisthopappus taihangensis*. For this study, potential major indicators were investigated to explore those associated with the genetic variations in *Opisthopappus* populations.

Results: The findings revealed that a significant genetic differentiation existed between *O. longilobus* and *O. taihangensis* populations. All identified haplotypes/studied populations were divided into two distinct groups. The considerable variability was due not only to geographic factors, but also environmental variables. At approximately 13.2 Ma of the Middle Miocene, *O. taihangensis* differentiated from *O. longilobus* under a strengthening East Asian monsoon and global cooling. Subsequently, with the uplift of the Taihang Mountains, the complex topography further promoted the variation between the two species and populations. Among environmental factors, five climatic variables, isothermality (bio3), min temperature of coldest month (bio6), mean temperature of wettest quarter (bio8), precipitation of driest month (bio14) and precipitation of wettest quarter (bio16) were demonstrated to be critical toward shaping the genetic components and genetic differentiation of the *Opisthopappus* species. Moreover, bio6 and bio14 had more driving force than others.

Conclusions: The temperature and precipitation of the Taihang Mountains during winter are the key environmental indicators responsible for the current genetic distribution. Our study revealed the importance of localized climatic events by taking the small-scale spatial effects of climate heterogeneity and its impacts on genetic variation into account.

Background

Understanding the processes that drive differentiation among populations and elucidating the mechanisms that underlie the origin and maintenance of genetic variation are major aims and fundamental tasks in evolutionary biology [1-5], and they are also core issues in conservation biology [6, 7]. Myriad factors can impact the evolution and genetic differentiation of plant populations, of which geological events and climate oscillations are suggested as critical drivers [8-10]. In terms of geological events, mountain uplift leads to complex topographies that can segregate large plant populations into multiple smaller sub-populations and enhance differentiation between species or populations through geographical isolation. Further, climatic oscillations can shift in the ranges of species, and result in novel environments with increased variability [11, 12]. To adapt to different environments, organisms will evolve corresponding phenotypic variations and genetic differentiation [13, 14].

During this process, intensifying climate change had left an indelible imprint on the composition and divergence of populations or species [15, 16], which further greatly influenced the distribution patterns and shaped the genetic structures of populations [5, 17]. In general, the geographical processes of mountainous regions may influence the genetic makeup of plant populations over large spatial scales, whereas ecological processes from climate change may influence the genetic structures of plant populations at small spatial scales [5, 18-20].

The Taihang Mountains, with a north-south orientation (36–40ºN, 112–115ºE), are a prominent natural boundary in Northern China [21, 22], which have a geological developmental history of more than 2.5 billion years with a typical platform type crustal structure spanning Mesoproterozoic to Paleozoic Eras. Their distinct geotectonic positioning has produced a unique geological and geomorphic landscape. The Southern Taihang Mountains have existed as a major boundary of neotectonic deformation, represented by the Yuntai landform [21-23]. The Northern Taihang Mountains, represented by Zhangshiyan, Cnagyan, and Linlv landforms, are higher than their southern counterparts with an average elevation of 1500 m [21, 24, 25].

Climatically, the southern region is home to a warm temperate semi-humid climate with a mean annual temperature of 12.7°C and mean annual precipitation of 606.4 mm, while the northern region has a temperate continental monsoon climate with a mean annual temperature of ~10°C and mean annual precipitation of 700 mm [21, 24]. The topographic complexity of the Taihang Mountains coupled with increasing climate variability have impacted many organisms. Thus, this region may be regarded as a distribution and diversity center for numerous genera [21, 24].

As an important perennial herb germplasm resource of Asteraceae, *Opisthopappus* (which has been listed as a 2nd class protected plant in China [26]) grows only on the steep slopes and cliffs of the Taihang Mountains that span Shanxi, Hebei, and Henan Provinces [26-28]. Being a diploid species (2n = 18) [29], comprised of *Opisthopappus taihangensis* and *Opisthopappus longilobus* [25, 29, 30], this genus is endemic in China and has significant ornamental and medicinal value [25, 30-33].

Between *O. taihangensis* and *O. longilobus*, several morphologica [30, 34, 35] and genetic variations [36-40] have been observed. Nevertheless, the interspecific hybridization of these two species have never been reported.

In previous research, we presented certain correlations between the genetic differentiation and geographical distance of *O. taihangensis* and *O. longilobus* [37, 38]. However, precisely how the climatic heterogeneity of the Taihang Mountains influences the genetic differentiation of *Opisthopappus* populations has not been addressed as yet. Consequently, we hypothesized that the differentiation between these two species would be a hierarchically comprehensive process that might be initially impacted by climate oscillations, subsequently by geographical topography, and finally by the climatic heterogeneity of the Taihang Mountains.

For the present study (according to the hypothesis above), the roles and influences of environmental variables on species and population differentiation were investigated via the combination of geographic and climatic data, using single nucleotide polymorphism (SNP) and insertion-deletion (InDel). The aims of this study were to: (i) analyze the genetic variations between species and among all studied populations; (ii) investigate the evolutionary processes and histories of *O. taihangensis* and *O. longilobus* species; (iii) estimate the effect of geographical and climatic variables; (iv) confirm the key environmental indicators that drive this differentiation. These results would establish a theoretical foundation for the governance, protection, and utilization of *Opisthopappus*, while providing relevant and useful data for the study of additional plant species in the Taihang Mountains.

Results

Genetic variation of *Opisthopappus* populations

Among the designed SNP primers, eight pairs of primers produced repeatable, clear, and stable bands. The total length of eight SNP combination segments was 1921 bp, which contained 1870 conservative sites and 51 polymorphic sites.

Based on these sequences, seventy-five haplotypes were identified (Table 1). Therein, 47 haplotypes (H1-H47) were detected in *O. longilobus* and 28 haplotypes (H48-H75) in *O. taihangensis*. No shared haplotypes were detected between *O. longilobus* and *O. taihangensis*. For *O. longilobus*, the H5 haplotype was the most widely distributed (shared by three populations), whereas the other 41 haplotypes were distributed only in a single population (Table 1, Fig. 1). For *O. taihangensis*, the H50 and H52 haplotypes were the most widely distributed (both shared by six populations), the remaining 15 haplotypes were detected only in a single population. The value of the genetic differentiation coefficient N_{ST} was 0.743, which was significantly larger than the value of G_{ST} = 0.105, indicated that there were significant phylogeographic structures in the *Opisthopappus* genus.

Both the SNP sequences and InDel data revealed a high genetic diversity in *Opisthopappus*. According to the SNP data, the haplotype diversity (*H*d) was 0.98571 and the total nucleotide polymorphism (π) was 0.00730. As relates to the InDel data, the Nei's gene diversity index (*H*), polymorphic loci ratio (*PPL*), and Shannon's polymorphism information index (*I*) were 0.2460, 93.85%, and 0.3887, respectively (Table 1). A relatively high level of genetic diversity was found in both *O. longilobus* and *O. taihangensis* (Table 1). Moreover, the genetic parameters of *O. longilobus* were higher than that of *O. taihangensis*.

Significant genetic variations occurred between *Opisthopappus* populations ($F_{ST\ SNP}$ =0.743, $G_{ST\ InDel}$ =0.6973), which was consistent with the AMOVA results. Based on the SNP sequences, 80% of the mutations occurred between *O. longilobus* and *O. taihangensis* (F_{CT} = 0.8003, P < 0.01), 15% of the molecular variations within the populations (F_{ST} = 0.8460, P < 0.01), and only 5% of the molecular

variations between populations within the groups (F_{SC} = 0.2287, P < 0.01). For the InDel data, the genetic variation distribution trend was similar to the SNP sequences. The results verified that molecular variations existed primarily between these two *Opisthopappus* species (Table 2).

A phylogenetic tree, which was constructed based on eight SNP combination fragments (Additional file 1: Figure S1A), revealed that all individuals were clearly divided into two groups. Group I contained all *O. longilobu*s individuals, while group II contained all *O. taihangensis* individuals. UPGMA cluster analysis (Additional file 1: Figure S1B) performed based on Nei's genetic distance revealed that twenty-four populations were separated into two (*O. longilobus* and *O. taihangensis*) clusters.

When $\triangle K$ (mean (|L"(K)|/sd(L(K))) attained a maximum value, K=2 was taken on both SNP and InDel data. The most significant possibilities were gathered into two groups (Fig. 2). For the DAPC analysis, the first ten principal components (PCs) represented a 91.1% variation of genetic components. *O. longilobus* populations were clustered into a single group, and *O. taihangensis* populations were clustered into another group (Fig. 3A). At the population level, obvious genetic structures were observed between different species, whereas the populations within species were not clearly distinguished (Fig. 3B). The results were confirmed by the K-W test (Additional file 2: Figure S2).

Historical dynamics of *Opisthopappus* populations

A Bayesian Inference tree was developed for seventy-five haplotypes (Fig. 4), which were segregated into two distinct branches. One branch contained the H1-H47 haplotypes from *O. longilobus*, whereas the other contained the H48-H75 haplotypes from *O. taihangensis*. Among the 28 haplotypes in *O. taihangensis*, the H71 haplotype from the GS population formed a separate sub-branch, which was consistent with the individual ML tree (Additional file 1: Figure S1A). Further, the haplotype network (Additional file 3: Figure S3) presented a similar haplotype distribution pattern with the haplotype phylogenetic tree.

At approximately 13.2 Ma in the Middle Miocene, *Opisthopappus* began to differentiate into two major lineages: *O. longilobus* and *O. taihangensis*. Within *O. taihangensis*, the differentiation time was at 8.4 Ma between haplotypes, while the haplotypes of *O. longilobus* were differentiated at ~9.4 Ma (Fig. 4). The recent differentiation time of intraspecific haplotypes for both two species were (e.g. H61 and H58 of *O. taihangensis* and H28 and H41 of *O. longilobus*) approximately in the Quaternary Era.

The neutral test and mismatch distribution analysis (MDA) (Additional file 4: Table S1) revealed that both *Opisthopappus* species and the genus had experienced a recent expansion based on significantly negative Fu's *Fs* (-25.2963, -18.5566, and -24.1000 for *O. longilobus*, *O. taihangensis*, and *Opisthopappus*, respectively, P < 0.05) values and non-significant SSD and Rag values (P > 0.05).

According to ABC analysis, scenario 3 was considered to be the most strongly supported evolutionary event within the three scenarios (Additional file 5: Table S2, Additional file 6: Figure S4), namely, *O. taihangensis* was a descendant of *O. longilobus*. The estimated effective population size of *O. longilobus* (5.63E+04, 95% CIs: 1.51E+04~2.10E+05) was smaller than that of *O. taihangensis* (2.13E+05, 95% CIs: 6.06E+04~5.31E+05).

The historical gene flow generated using MIGRATE were low between the two species, $Nm_{O.\ longilobus \rightarrow O.\ taihangensis}$ = 0.3813 (95% CIs: 0-1.8790) and $Nm_{O.\ taihangensis \rightarrow O.\ longilobus}$ = 0.7860 (95% CIs: 0~3.1748). The mean contemporary gene flow (m) between the two species was also low. The migration occurred from *O. longilobus* to *O. taihangensis*, and the reverse was 0.0059 (95% CIs: 0~0.0173) and 0.0052 (95% CIs: 0~0.0156), respectively. However, the migration rates within each species were relatively high (*O. longilobus*: 0.9941 (95% CIs:0.9827~1.0055), *O. taihangensis*: 0.9948 (95% CIs: 0.9844~1.0052)) as estimated by BAYESASS.

Environmental heterogenicity driven on *O. longilobus* and *O. taihangensis*

Between the nineteen extracted bioclimatic variables (Bioclim), following the removal of 14 bioclimatic variables of collinearity, the remaining five bioclimatic variables were: bio3, bio6, bio8, bio14, and bio16 (Additional file 7: Table S3, Additional file 8: Figure S5). One-way ANOVA revealed that all the remaining variables distributed along the two species were significantly different, except for bio8 (P = 0.8232, Additional file 7: Table S3).

The partial correlation of bioclimatic variables via PCA revealed that the explanatory direction was different for these variables (Fig. 5). The PCA plot drawn on the first two axes explained 48.39% and 33.79% of the variation in the climate variables, respectively. Although climate distribution conditions are similar in some populations, they do not overlap completely.

To assess whether geographic or environmental differences might drive genetic divergence, both Mantel and partial Mantel tests were conducted. They revealed that there were non-significant associations between geographical and environmental distances (r = 0.0161, P = 0.421). Nevertheless, significant correlations were discovered in both the genetic and geographic distances (r = 0.5039, P = 0.001), and genetic and environmental distances (r = 0.1512, P = 0.024) (Table 3). The partial Mantel tests also detected significant correlations between genetic and environmental distances conditioned on the geographic effect (r = 0.1656, P = 0.011)

Similar results, which revealed that geographic and environmental distances affected genetic distances, were obtained by MMRR analyses (Table 4). The effects were that both the geographic (coefficient = 0.0029, r^2 = 0.2539, P = 0.001) and environmental (coefficient = 0.0038, r^2 = 0.0228, P = 0.016) distances were significantly related to the genetic distance. The joint effects of both the geographical and environmental distances also impacted the genetic distance significantly (r^2 = 0.0404, coefficient $_{geo}$ =

0.0029, P = 0.001, coefficient $_{env}$ = 0.0036, P = 0.019). The above results indicated that the genetic differentiation of populations was significantly and linearly correlated with geographic and/or climatic differentiation.

The full RDA model, including geographic distribution and climatic factors could explain 72.75% (conditioned: 11.11%, constrained: 61.64%, Table 4) of the variation of the genetic components. The partial RDA, while conditioned on the geographic distribution (coordinates) of sites, found a significant effect of climate variables following the removal of the effects of isolation by distance (Proportion = 61.64%, adj R^2 = 0.5099, P = 0.01). The ANOVA further indicated that bio6 and bio14 significantly explained the genetic components (P < 0.05) of the population, with high explanatory proportions (25.03% and 11.92%, respectively, Table 4).

The distribution of climatic variables along the ordination axis was further examined by GLM (Table 4). Three bioclimatic variables, bio6 (adj R^2 = 0.4768, F = 21.96, P < 0.05), bio14 (adj R^2 = 0.2047, F = 6.92, P < 0.05), and bio16 (adj R^2 = 0.1535, F = 5.17, P < 0.05), had significant F statistics. Four variables (except bio8) correlated significantly with the ordination axis1 of dbRDA, while only bio14 significantly correlated with axis2.

The high adjusted R^2 indicated that these were sufficiently explained by the two dbRDA axes. Consistent estimates of influence were obtained even when the order of predictors was altered, thus, these variables were highly relevant for explaining the genetic structure of the population. However, the adjusted R^2 of bio8 might be too small to be meaningful in either dbRDA or GLM.

Finally, the importance of climatic factors in the RDA model was accessed by the R package ‘vegan’ and ‘packfor’, respectively (Table 5). For the former, bio6 was selected out in step one according to the P and AIC values. Subsequently, bio14 was selected out in step two, and there was no significant climate factor for further steps. For the latter, the adj R^2 was relatively high for reserved explanatory variables (0.4304 and 0.1052 for bio6 and bio14, respectively). Both methods considered that bio6 and bio14 could be the most important climate variables that contributed to the genetic differentiation with strong explanatory capabilities.

The scatter and ordisurf plots (Fig. 6) for the most important variables (bio6 and bio14) congruently showed broader ranges of environmental contours in *O. taihangensis* populations compared with *O. longilobus* populations. The populations of the two species exhibited a significantly different distribution of the dbRDA space along axis 1, but not axis 2, which were similar to the DAPC clustering patterns. The distribution of populations within species were too close to be distinguished. Ordisurf plots clearly illustrated the climatic differentiation between *O. longilobus* and *O. taihangensis*.

Discussion

Genetic differentiation results from geographic topology and climate heterogeneity

High genetic diversity in populations of organisms is a successful strategy that facilitates their adapting to various habitats and environmental conditions [41, 42] toward increasing their effective colonization in new habitats [42-44]. In this study, high genetic diversity was observed in *O. longilobus* or *O. taihangensis* (Table 1), which was consistent with previous research by ISSR [36] and SRAP [38] markers. This was similar to other endemic and endangered species, such as *Helianthemum songaricum* [45], *Sinowilsonia henryi* [46], and *Taihangia rupestris* [47]. Meanwhile, *O. longilobus*, as an ancestor of *O. taihangensis*, might have a relatively higher genetic diversity than *O. taihangensis* (Table 2, Additional file 5: Table S2, Additional file 6: Figure S4) [48].

The overall genetic variation existed primarily between populations of different species, which was revealed by AMOVA (Table 2) and the K-W test (Additional file 2: Figure S2). In phylogenetic analysis, the populations formed two separated clusters (Additional file 1: Figure S1), and the identified haplotypes were divided into two groups (Additional file 3: Figure S3), which was confirmed by STUCTURE (Fig. 2) and DAPC (Fig. 3) analyses. The two detached clusters or groups were identified as *O. longilobus* and *O. taihangensis* species. These analyses proposed that the speciation of *O. longilobus* and *O. taihangensis* was complete, or at least that species divergence was significantly reflected in the overall genetic differentiation.

The genetic differentiation of *Opisthopappus* was a shared outcome contributed by both geography and environment. Genetic variation proportioning between populations reflected the consequences of the interactions of multiple factors, while population divergence was driven via isolation by distance (IBD) and/or isolation by environment (IBE) [49-51]. Of the 26 studies related to plants summarized by Sexton et al. (2014) [52], 38.5% found both IBD and IBE patterns, and 30.8 and 11.5% found IBD and IBE, respectively. The IBD and IBE patterns in this study were also concurrently confirmed by Mantel tests and MMRR analysis (Table 3). The gene flow between populations could follow both the IBD and IBE patterns, which corresponded to geographic distance and environmental context [52]. The contemporary gene flow between two species estimated by BAYESASS was fairly low, and the migration mainly occurred within the species.

With the rapid uplift of the Taihang Mountains, its complex topographies began to serve as a significant geographical barrier in Northern China [22], which could segment *O. taihangensis* and *O. longilobus* into spatially isolated subpopulations following primary differentiation. The many gullies and gorges of the Taihang Mountains likely prevented pollen exchange between populations (being a wind-pollinator species). With the decline in genetic exchange, population divergence would increase [52, 53]. Moreover, the environments (including temperature and precipitation) for each population was diversified due to the variable localized climates of the Taihang Mountains. This situation enabled these isolated subpopulations to eventually evolve into genetically distinct lineages via the accumulation of genetic differences and adaptations to local environments [16, 17, 54]. The adaptive divergence between populations was typically initiated by single or simple environmental differences, followed by the expansion of accumulating genetic divergence [5, 55-57]. The ANOVA (Additional file 7: Table S3) and PCA (Fig. 5) results revealed that reversed climate factors initiated a significantly different distribution

between the two species. Environmental heterogeneity caused significant genetic divergence (Fig. 3, Additional file 2: Figure S2) and contributed to a high proportion of genetic variation in the populations (Table 2) [58, 59]. The partial dbRDA also indicated that ~60% of the variation between populations could be explained by climatic heterogeneity (Table 4).

Evolutionary dynamics associated with geological history and climate oscillations

In the Middle Miocene (about 13.2 Ma), the *Opisthopappus* haplotype began to diverge between *O. longilobus* and *O. taihangensis*. During the Miocene period, the Taihang Mountains were still not formed [21, 25]. About 8-15 million years ago, the rapid uplift of the Qinghai-Tibet Plateau occurred, which lead to the monsoon climate in East Asia [60, 61]. Meanwhile, the global climate began to cool during/following the Tertiary Period [20, 62], and the East Asian monsoon climate increased dramatically. The strengthening of the East Asian monsoon climate and global cooling might have driven the initial differentiation between *O. longilobus* and *O. taihangensis*. Subsequently, with changing monsoon patterns in East Asia during the Late Miocene, intraspecies differentiation ensued, with the *O. taihangensis* haplotype differentiating at ~8.4 Ma, and that of *O. longilobus* at ~9.4 Ma (Fig. 4). Climate fluctuations would further promote differentiation between haplotypes or species.Additionally, the most recent diversification of haplotypes within species occurred primarily during the Quaternary glacial period (Fig. 4), which had a considerable impact on the distribution of modern animals and plants [16, 63-65]. Glacial and interglacial cycles played significant roles in the formation and differentiation of species. Many plants transitioned to refuges during glacial periods, while undergoing migration, repeated contraction, and expansion during the interglacial periods [17, 20, 66, 67].

Based on the results of mismatch distribution analysis and neutrality tests, both *O. taihangensis* and *O. longilobus* showed signs of expansion (Additional file 4: Table S1). This was also confirmed by the *Opisthopappus* haplotype distribution and network (Fig. 1, Additional file 3: Figure S3). During the Quaternary, the paleovegetation of the Taihang Mountains repeatedly appeared as replacement species during grassland and forest cycles [68]. Emerging grasslands might have served as a transitional corridor that provided opportunities for populations to expand and colonize.

Temperature and precipitation during winter are key environmental factors responsible for the current genetic distribution

Temperature and precipitation were commonly found to play prominent roles as selective drivers for variations in various plant species [17, 57, 65, 69]. Significant temperature fluctuations might contribute to the physiological states, metabolic levels, and genetic alterations of plants, to further drive the genetic makeup of the population. Precipitation during seed germination and growth could affect the demographic size as well as influencing successful seed colonization. The distribution areas of

Opisthopappus belong to warm temperate continental monsoon and temperate continental semi-humid monsoon climates [21, 25]. Various sample sites had different climate conditions. For example, Jiaozuo and Xinxiang City are located in a warm temperate continental monsoon climate zone, with an average annual temperature of 14°C. The hottest temperatures occur during July with an average of 27°C, the coldest in January with a mean of 0.1°C, with an annual average precipitation of 650 mm. In contrast, Lingchuan City in Shanxi Province belongs to a temperate continental semi-humid monsoon climate. The average annual temperature is ~7.9°C, with the most extreme minimum temperature on record being -26.2°C, and an annual average precipitation of from 600 mm to 700 mm [24]. These localized climate differentiations may cause divergent adaptability of *Opisthopappus* populations.

In this study, bio6 and bio14 (Table 4) were considered to be the most powerful factors through the forward selection of dbRDA (Table 5). For bio6, the minimum temperature of the coldest month belongs to the temperature dimension, while for bio14, the precipitation of the driest month belongs to the precipitation dimension. According to the WorldClim database, January was both the coldest and the driest month (average temperature -2.7°C, average precipitation 4.375mm) (Additional file 9: Figure S6). In our field investigation, the populations of *Opisthopappus* were sensitive to the climate conditions in January. During this month, *Opisthopappus* individuals required a certain temperature and precipitation to meet the demands of their growth, which indicated that *O. taihangensis* and *O. longilobus* could grow under mild winter conditions. Therefore, bio6 and bio14 were considered to be the key climate factors for governing both the genetic distribution and genetic structure of *Opisthopappus.*

Conclusions

Spatial environmental heterogeneity is typically suggested as a critical driver that leads to population differentiation, and even the acceleration of speciation. Here, we provided comprehensive evidence, including genetics, geographical conditions, climate variables, and evolutionary processes to validate the differentiation of two *Opisthopappus* species. When all of these aspects were considered, *O. taihangensis* and *O. longilobus* were clearly distinct. At ~13.2 Ma in the Miocene, *O. taihangensis* was derived from *O. longilobus*, which was the result of an East Asian monsoon and a globally cooling climate. With the uplift of the Taihang Mountains and the glacial–interglacial cycles during Quaternary period, *Opisthopappus* was isolated into multi-segregated populations. Subsequently, the complex topography, local climate, and environment of the Taihang Mountains began to affect the genetic structures and distribution patterns of the two species. Among five reserved environmental variables, bio6 (Min Temperature of Coldest Month) and bio14 (Precipitation of Driest Month) were considered to be two most important variables. The temperature and precipitation of the Taihang Mountains during winter are the key environmental indicators responsible for the current genetic distribution. Our study revealed the importance of localized climatic events by taking the small-scale spatial effects of climate heterogeneity and its impacts on genetic variation into account.

Methods

Sample collection

Our study was conducted in accordance with the laws of the People's Republic of China, and field collection was approved by Chinese Government. All researchers received permission letters from the College of Life Science, Shanxi Normal University, to collect the samples. The samples were taxonomically identified based on their phenotype by Junxia Su (Associate Professor of systematic botany) at the Shanxi Normal University. The voucher specimens were deposited in the herbarium of College of Life Science, Shanxi Normal University (No:20170105030-20170105050).

Eleven populations of *O. longilobus* and thirteen populations of *O. taihangensis* were sampled, which covered the distribution ranges of *Opisthopappus* (Additional file 10: Table S4, Fig.1). Individuals growing at a common site were regarded as a single "population". Fresh young leave devoid of disease or insect pests were selected for each of the sample sites, where 10-15 individuals from each population were collected. The collected samples were placed into sealed bags filled with silica gel, dehydrated/quickly dried, and stored at 20°C for later use. A global positioning system (GPS) was employed to demarcate each sample site and record the longitude, latitude, and altitude of each population.

PCR amplification, sequencing and genotyping

The total genomic DNA was extracted using the modified 2 × CTAB method [70]. The quality of DNA was measured using an ultraviolet spectrophotometer and 0.8 % agarose gel electrophoresis and stored at -20°C for further use. The SNP and InDel primers were designed according to the predicted locus from the transcriptome sequences of *Opisthopappus* (Additional file 11: Table S5).

For the SNP primers, the 20 μL PCR reaction contained 10 μL 2 × MasterMix, 2 μL template DNA (30 ng/μL), 1 μL primer S (10 μM), 1 μL primer A (10 μM), and 6 μL ddH_2O. The PCR procedure proceeded as follows: pre-denaturation at 94°C for 5 min., denaturation at 94°C for 1 min, annealing temperature based on each primer (Table 2) setting for 1 min, elongation at 72°C for 1.5 min., repeated for 35 cycles, last elongation at 72°C for 10 min, and preservation at 4°C. The PCR products detected using 2% agarose gel electrophoresis were confirmed via an automatic analysis electrophoresis gel imaging system, then sent to Sangon Biotech (Shanghai) for sequencing.

For the InDel primers, the PCR reaction was 20 μL, which contained 10 μL 2 × MasterMix, 3 μL template DNA (30 ng/μL), 1 μL primer S (10 μM), 1 μL primer A (10 μM), and 5 μL ddH_2O. The PCR procedure was as follows: pre-denaturation at 94°C for 1 min, denaturation at 94°C for 1 min, annealing temperature based on each primer (Additional file 11: Table S5) setting for 1 min, elongation at 72°C for 1 min, repeated for 35 cycles, last elongation at 72°C for 10 min, preservation at 4°C. The PCR products were

detected using 8% polyacrylamide gel electrophoresis. The presence or absence of each InDel fragment were coded as ‘1’and ‘0’ respectively.

Population genetic differentiation analyses

For the SNP sequences, haplotypes, haplotype frequencies, haplotype diversity (*H*d), and nucleotide diversity (π) were calculated using DNASP 5.10 [71]. The genetic differentiation parameters G_{ST} and N_{ST} were examined by PERMUT 2.0 [72] based on the haplotype frequency. For the InDel data, the genetic characteristics, Nei's gene diversity index (*H*), Shannon’s information index (*I*), and the percentage of polymorphic loci (*PPL*), were calculated by POPGENE 1.31 [73].

An analysis of molecular variance (AMOVA) was implemented by ARLEQUIN 3.5 [74] and GENALEX 6.5 [75] to detect the distribution of genetic variations within and between populations or species. Subsequently, the values of F_{ST}, $F_{CT,}$ and F_{SC} [76] were calculated based on hierarchical AMOVA, and the permutation test was set to 1000.

Cluster analysis based on the maximum likelihood (ML) method and Nei’s genetic distance, respectively, was performed using MEGA 7.0 [77]. Bayesian clustering analysis (BCA) was employed to examine the similarity and divergence of genetic components between populations, and performed using STRUCTURE 2.2 [78] for both SNP sequencing and InDel data. The posterior probability of grouping number (K =2–24) was estimated through 10 independent runs using 500,000 step Markov chain Monte Carlo (MCMC) replicates, following a 1,000,000-step burn-in for each run to evaluate consistency.

The best grouping number was evaluated by $\triangle K$ [79] in STRUCTURE HARVESTER 0.6.94 [80]. These 10 runs were aligned and summarized using CLUMPP 1.1.2 [81] and the visualization of the results was plotted using DISTRUCT 1.1 [82].

To test the genetic divergence between populations or species, a discriminant analysis of principal components (DAPC) was implemented by the function dapc in the R package ‘adegenet’ [83], which initially transformed the genetic data using principal components analysis (PCA) results, and subsequently performed a discriminant analysis on the retained principal components.

The properties of the “without *a priori*”, using partial synthetic variables to minimize variations within groups [84], might assist with objectively evaluating the artificial classification of *O. taihangensis* and *O. longilobus*. Kruskal-Wallis tests for the first two principal components (PCs), and the first two linear discriminants (LDs) of DAPC, were conducted to examine the genetic divergence between populations and species.

Population demographic history inferring

A network relationship was generated through the median-joining method in Network 5.0 [85], to investigate the evolutionary relationships between the *Opisthopappus* haplotypes. BEAST 1.84 [86] was employed to estimate the differentiation and diversification time between haplotypes, and *Chrysanthemum indicum* was selected as an outgroup. JMODELTEST 1.0 [87] was utilized to calculate the best nucleotide substitution model, and GTR was finally confirmed as the optimal model. An uncorrelated lognormal relaxed molecular clock model was adopted. Markov chain Monte Carlo (MCMC) was repeated 8×10^7 times by sampling every 80,000 generations. Since the rate of base substitution in *Opisthopappus* has not been calculated as yet, we estimated the differentiation time by using the reported average value 2×10^{-9} of the angiosperm genome. TRACER 1.5 [86] was used to check the convergence of the framework, which ensured that every tested parameter was greater than 200.

To assess whether the species had experienced significant expansion, we used ARLEQUIN 3.5 [74] to calculate the Tajima's *D* [88] and Fu's F_S [89] values. Moreover, the sum of square deviation (SSD) and raggedness index (Rag) in the mismatch distribution analysis (MDA) was also performed in ARLEQUIN 3.5 [74]. The process employed a 1000 step permutation test.

Approximate Bayesian computation (ABC) analysis, provided by DIY-ABC 2.1.0 [90], enabled the estimation of complex evolutionary population histories. Based on the estimated genetic variations, genetic structures, and current geographic distributions, three evolutionary scenarios were proposed. Scenario 1: *O. longilobus* and *O. taihangensis* were differentiated from a common ancestral population during the same period. Scenario 2: *O. taihangensis* was an ancestral population, and the *O. longilobus*was differentiated from *O. taihangensis*. Scenario 3: *O. longilobus* was the ancestral population, and *O. taihangensis* was differentiated from *O. longilobus*. Each scenario was performed with 1,000,000 simulations. There was up to 95% confidence intervals through logistic regression, and the posterior probability of each evolutionary event was estimated using the observation data closest to 1% of the simulated data. An optimal evolutionary scenario was selected according to the maximum posterior probability value, and the parameters including effective population size and divergence generation was estimated under the optimal scenario.

In addition, the historical and contemporary gene flows were estimated within the two *Opisthopappus* species by MIGRATE-N 3.6 [91] and BAYESASS 3.0 [92]. In MIGRATE-N 3.6, maximum-likelihood analyses were performed using 10 short chains (10^4 trees) and three long chains (10^5 trees) with 10^4 trees discarded as an initial burn-in' and astatic heating scheme at four temperatures (1, 1.5, 3, and 1000,000). To ensure the consistency of estimates, we repeated this procedure five times and reported average maximum-likelihood estimates with 95% confidence intervals. The parameters θ and *M* were estimated using a Bayesian method, which could be employed to estimate the number of migrants per generation (*N*m) into each population using the equation $4Nm = \theta * M$. When estimating the contemporary gene flow using BAYESASS 3.0, the parameters were examined including migration rates (m), allele frequencies (a) and inbreeding coefficients (f) to ensure that the optimal acceptance rates of the three parameters fell within the range of 20-60%. Ten independent runs were executed to minimize the convergence problem.

The result with the lowest deviance was adopted according to the method of Meirmans [93], where the 95% credible interval was estimated as m ± 1.96 x standard deviation (SD).

Environmental variables influence analyses

Nineteen bioclimatic variables (Bioclim) representing Grinnellian niches [94], which are defined as the scenopoetic environmental variables of a species required to survive, were downloaded from the WorldClim database (http://www.worldclim.org/) with a resolution of 30 arc-sec (~1×1 km) and extracted using the R package 'raster'. Subsequently, a correlation analysis of the bioclimatic variables was performed prior to further analysis to reduce multicollinearity. The variance inflation factor (VIF) was computed for climate variables using the "vif" function in the R package 'usdm' [95]. Climate variables with VIF > 20 [96], and those highly correlated with other variables ($|r| > 0.8$) were removed. Finally, five climate variables were retained, namely bio3 (Isothermality), bio6 (Min Temperature of Coldest Month), bio8 (Mean Temperature of Wettest Quarter), bio14 (Precipitation of Driest Month), and bio16 (Precipitation of Wettest Quarter) (Additional file 7: Table S3, Additional file 8: Figure S5).

The statistics and distribution of reserved climate variables were displayed using box-plots grouped by the two species (Additional file 8: Figure S5) and the significance was tested by one-way ANOVA (Additional file 7: Table S3). The principal component analysis (PCA) of independent climatic variables to reduce the dimensionality that defined the niche space allowed for the comparison of the integrity of climate variables between *O. longilobus* and *O. taihangensis*.

To test how the geographical distance and environmental differences impacted genetic differentiation, the Mantel test of Pearson correlation was performed between genetic, geographic, and environmental (climatic) distances. Pairwise F_{ST} distance calculated in ARLEQUIN 3.5 [74] was used as genetic distance. The geographic distance was estimated using the GENALEX 6.5 [75] according to three-dimensional factors (latitude, longitude, and altitude). The environmental distance was calculated using the Euclidean distance with PASSAGE 2.0 [97].

A partial Mantel test between genetic, environmental, and geographic distances was conducted under the condition of the geographic or environmental effect, respectively. Further, a multiple matrix regression with randomization (MMRR) was performed to test whether the genetic distance responded to variations in the geographic and/or environmental distances. The Mantel test was performed in the R package 'vegan' [98], and MMRR analysis was performed using the R package 'PopGenReport' [99]. Regression coefficients of the Mantel test (r) and MMRR (r^2) and their significance were determined based on 9,999 random permutations.

Distance based redundancy analyses (dbRDA) were performed to elucidate whether the climatic variables conditioned on the geographic distribution explained the genetic differentiation of the populations using the R package 'vegan' [98]. Firstly, a distance-based principal coordinate analysis (PCoA) of the genetic

data at the species level was performed to generate several principal coordinates (PCs) using the R package 'ape' [100]. Next, five retained climatic variables were employed as explanatory variables conditioned on geographic factors, and permutation tests for each climatic variable were performed using the "anova. cca" [101] function in the R package 'vegan'.

The distribution of climate variables along the ordination axes was further analyzed using a generalized linear model (GLM). Finally, two different methods were employed to select the key factors which might be essential for driving the variation of *Opisthopappus* populations: (i) by the Akaike Information Criterion (AIC) value and the significance level P value of *F*-statistic [102] using the "ordistep" function of R package 'vegan'. During each selecting step, only one factor was retained with the most significant P value. If the P value of the two variables was equal, the variable with the smaller AIC value was selected. (ii) by the double stopping criterion [103] using the "forward.sel" function of R package 'packfor' [104]. Significance was determined using 999 permutations in the RDA analysis. The first two RDA axes and the selected key factors were employed to construct the ordination and ordisurf plots of the dbRDA.

Abbreviations

bio3: isothermality; bio6: min temperature of coldest month; bio8: mean temperature of wettest quarter; bio14: precipitation of driest month; bio16: precipitation of wettest quarter; AMOVA: analysis of molecular variance; BCA: Bayesian clustering analysis; PCA: principal components analysis; PC: principal components; LD: linear discriminants; K-W test: Kruskal-Wallis test. MDA: mismatch distribution analysis; ABC: approximate Bayesian computation; MMRR: multiple matrix regression with randomization; dbRDA: distance based redundancy analysis; GLM: generalized linear model; AIC: Akaike information criterion.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files. The datasets used and/or analyzed during the current study are available from the authors on reasonable request (Hang Ye, SXSDyehang@hotmail.com; Yiling Wang, ylwangbj@hotmail.com).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was funded by the National Natural Science Foundation of China (31970358 to Y.W.). The funding agency did not play a role in the experimental design, results analysis or writing of the manuscript, but did provide financial support for the manuscript.

Authors' contributions

HY analyzed the data and wrote the first draft. YW proposed the main structure of this study. WR provided useful advice. GS revised the text. ZW and HH revised the figures. JW, YG and WH collected the sample cores. All authors contributed to the final manuscript.

Acknowledgements

The authors thank Junxia Su for her assistance with the sample identification.

References

1. Mayr E. Animal species and evolution. The Belknap Press of Harvard University Press, Cambridgem Massachusetts. 1963.
2. Knowles LL. Tests of pleistocene speciation in montane grasshoppers (genus *melanoplus*) from the sky islands of western north america. Evolution. 2000;54:1337-48.
3. Callahan CM, Rowe CA, Ryel RJ, Shaw JD, Madritch MD, Mock KE. Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). J Biogeogr. 2013;40:1780-91.
4. Carvalho CdS, Ballesteros-Mejia L, Ribeiro MC, Côrtes MC, Santos AS, Collevatti RG. Climatic stability and contemporary human impacts affect the genetic diversity and conservation status of a tropical palm in the Atlantic Forest of Brazil. Conserv Genet. 2017; 18:467-78.

5. Ben-Menni Schuler S, López-Pujol J, Blanca G, Vilatersana R, Garcia-Jacas N, Suárez-Santiago VN. Influence of the Quaternary glacial cycles and the mountains on the reticulations in the subsection *Willkommia* of the genus *Centaurea*. Front Plant Sci. 2019;10:303.
6. Lavergne S, Mouquet N, Thuiller W, Ronce O. Biodiversity and climate change: integrating evolutionary and ecological responses of species and communities. Annu Rev Ecol Evol S. 2010;41:321-50.
7. Mouquet N, Devictor V, Meynard CN, Munoz F, Bersier L-F, Chave J, Couteron P, Dalecky A, Fontaine C, Gravel D et al. Ecophylogenetics: advances and perspectives. Biol Rev. 2012; 87:769-85.
8. Young A, Boyle T, Brown T. The population genetic consequences of habitat fragmentation for plants. Trends Ecol Evol. 1996;11:413-8.
9. Hewitt G. The genetic legacy of the Quaternary ice ages. Nature. 2000;405:907-13.
10. Xie DF, Li MJ, Tan JB, Price M, Xiao QY, Zhou SD, Yu Y, He XJ. Phylogeography and genetic effects of habitat fragmentation on endemic *Urophysa* (Ranunculaceae) in Yungui Plateau and adjacent regions. PLoS One. 2017;12:e0186378.
11. Drake JM. Population effects of increased climate variation. P Roy Soc B-Biol Sci. 2005; 272:1823-7.
12. Stojanova B, Šurinová M, Klápště J, Koláříková V, Hadincová V, Münzbergová Z. Adaptive differentiation of *Festuca rubra* along a climate gradient revealed by molecular markers and quantitative traits. PLoS One. 2018;13:e0194670.
13. Sexton JP, McIntyre PJ, Angert AL, Rice KJ. Evolution and ecology of species range limits. Annu Rev Ecol Evol S. 2009;40:415-36.
14. Bernatchez L. On the maintenance of genetic variation and adaptation to environmental change: considerations from population genomics in fishes. J Fish Biol. 2016;89:2519-56.
15. Hewitt GM. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc. 2008;58:247-76.
16. Deli T, Kiel C, Schubart CD. Phylogeographic and evolutionary history analyses of the warty crab Eriphia verrucosa (Decapoda, Brachyura, Eriphiidae) unveil genetic imprints of a late Pleistocene vicariant event across the Gibraltar Strait, erased by postglacial expansion and admixture among refugial lineages. BMC Evol Biol. 2019;19:105.
17. Contreras-Moreira B, Serrano-Notivoli R, Mohammed NE, Cantalapiedra CP, Beguería S, Casas AM, Igartua E. Genetic association with high-resolution climate data reveals selection footprints in the genomes of barley landraces across the Iberian Peninsula. Mol Ecol. 2019; 28:1994-2012.
18. Sacks BN, Brown SK, Ernest HB. Population structure of *California coyotes* corresponds to habitat-specific breaks and illuminates species history. Mol Ecol. 2004;13:1265-75.
19. He SL, Wang YS, Li DZ, Yi TS. Environmental and Historical Determinants of Patterns of Genetic Differentiation in Wild Soybean (*Glycine soja* Sieb. et Zucc). Sci Rep. 2016;6:22795.
20. Ye JW, Zhang ZK, Wang HF, Bao L, Ge JP. Phylogeography of *Schisandra chinensis* (Magnoliaceae) reveal multiple refugia with ample gene flow in Northeast China. Front Plant Sci. 2019;10:199.

21. Gong M. Uplifting process of southern Taihang Mountain in Cenozoic. Chinese Academy of Geological Science. Thesis for Doctor'Degree. 2010.
22. Zhang Y, Ma Y, Yang N, Shi W, Dong S. Cenozoic extensional stress evolution in North China. J Geodyn. 2003;36:591-613.
23. Zhang M, Li P. Discussion on the main uplift period of the Southern segment of Taihang Mountains. Territory & Natural Resources Study. 2014;4:55-7.
24. Zhu L. Spider community structure in fragmented habitats of Taihang Mountain area, China. Thesis for Master' Degree, Hebei University. 2008.
25. Wu C, Zhang X, Ma Y. The Taihang and Yan mountains rose mainly in Quaternary. Norht China Earthquake Sciences. 1999;17:1-7.
26. Zhao HB, Chen FD, Chen SM, Wu GS, Guo WM. Molecular phylogeny of *Chrysanthemum*, Ajania and its allies (Anthemideae, Asteraceae) as inferred from nuclear ribosomal ITS and chloroplast *trn*L-F IGS sequences. Plant Syst Evol. 2010;284:153-69.
27. Yang D, Hu X, Liu Z, Zhao H. Intergeneric hybridizations between *Opisthopappus taihangensis* and *Chrysanthemum lavandulifolium*. Sci Hortic. 2010;125:718-23.
28. Tang F, Wang H, Chen S, Chen F, Teng N, Liu Z. First intergeneric hybrids within the tribe Anthemideae Cass. III. *Chrysanthemum indicum* L. Des Moul. × *Opisthopappus taihangensis* (Ling) Shih. Biochem Syst Ecol. 2012;43:87-92.
29. Hu X. Preliminary studies on inter-generic hybridization within *Chrysanthemum* in broad sense (III). Thesis for Master' Degree, Beijing Forestry University. 2008.
30. Ding BZ, Wang SY. Flora of Henan. Henan Science & Technology Press, Zhengzhou. 1998.
31. Zhang XW, Wei DW, Sun WY, Ye YZ. A Comparative study on antioxidant activity of alcoholic extracts of different organs of *Opisthopappus taihangensis* and *Dendranthema indicum*. Natural Product Research and Development. 2014;26:1120-6.
32. Yu K, Zhang XL, Zhao T, Ma PJ, Xu J. Inhibition of *Opisthopappus taihangensis* extracts on Hepatitis B virus. Drug Evaluation Research. 2016;39:61-5.
33. Wei DW, Xu MM, Sun WY, Jia CY, Zhang XW. Antioxidant activity of aqueous extracts from different organs of *Opisthopappus* Shih. Journal of Chinese Institute of Food Science and Technology. 2015;15:56-63.
34. Wu ZY. Compositae. Flora of China. Science Press, Beijing. 1993.
35. Jia R, Wang Y. Leaves Micromorphological characteristics of *Opisthopappus taihangensis* and *Opisthopappus longilobus* from Taihang Mountain, China. Vegetos. 2015;28:82-9.
36. Guo R, Zhou L, Zhao H, Chen F. High genetic diversity and insignificant interspecific differentiation in *Opisthopappus* Shih, an endangered cliff genus endemic to the Taihang Mountains of China. The Sci World J. 2013;2013:275753.
37. Wang Y. Chloroplast microsatellite diversity of *Opisthopappus* Shih (Asteraceae) endemic to China. Plant Syst Evol. 2013;299:1849-58.

38. Wang Y, Yan G. Genetic diversity and population structure of *Opisthopappus longilobus* and *Opisthopappus taihangensis* (Asteraceae) in China determined using sequence related amplified polymorphism markers. Biochem Syst Ecol. 2013;49:115-24.
39. Wang Y, Yan G. Molecular Phylogeography and population genetic structure of *O. longilobus* and *O. taihangensis* (*Opisthopappus*) on the Taihang Mountains. PLoS One. 2014;9:e104773.
40. Wang Y, Zhang C, Lin L, Yuan L. ITS sequence analysis of *Opisthopappus taihangensis* and *O. longilobus*. Acta Horticulturae Sinica. 2015;42:86-94.
41. Dasgupta N, Nandy P, Sengupta C, Das S. Genetic variation in relation to adaptability of three mangrove species from the Indian Sundarbans assessed with RAPD and ISSR markers. J For Res. 2018;29:301-10.
42. Nguyen XV, Jutta P. Assessment by microsatellite analysis of genetic diversity and population structure of Enhalus acoroides from the coast of Khanh Hoa Province, Vietnam. Acta Oceanol Sin. 2019;38:144-50.
43. Crawford KM, Whitney KD. Population genetic diversity influences colonization success. Mol Ecol. 2010;19:1253-63.
44. Drummond EBM, Vellend M. Genotypic diversity effects on the performance of *Taraxacum officinale* populations increase with time and environmental favorability. PLoS One. 2012; 7:e30314.
45. Han BC, Wei L, Yang XY, Liang XH, Wulanbateer S, S.D. Study on genetic diversity and phylogeographic structure of endangered *Helianthemum soongoricum*. Journal of Inner Mongolia Forestry Science and Technology. 2017;43:16-9.
46. He Y, Ji W, Zhou L, Ma J, Ji X, Li H. Genetic diversity and relationships of endangered *Sinowilsonia henryi* Hemsl by ISSR. Bulletin of Botanical Research. 2014;34:792-7.
47. Wang H, Cheng Y, Fang X, Ye Y, Zong Y, Cao R. Genetic diversity and differentiation in rare herb species *Taihangia rupestris* (Rosaceae). Acta Bot Boreali-Occidential Sinica. 2011;31:45-51.
48. Geng Q, Sun L, Zhang P, Wang Z, Qiu Y, Liu H, Lian C. Understanding population structure and historical demography of *Litsea auriculata* (Lauraceae), an endangered species in east China. Sci Rep. 2017;7:17343.
49. Lenormand T. Gene flow and the limits to natural selection. Trends Ecol Evol. 2002;17:183-9.
50. Shih KM, Chang CT, Chung JD, Chiang YC, Hwang SY. Adaptive genetic divergence despite significant isolation-by-distance in populations of taiwan cow-tail fir (*Keteleeria davidiana* Var. *formosana*). Front Plant Sci. 2018;9:92.
51. Endler JA. Gene flow and population differentiation. Science. 1973;179:243-50.
52. Sexton JP, Hangartner SB, Hoffmann AA. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution. 2014;68:1-15.
53. Star B, Spencer HG. Effects of genetic drift and gene flow on the selective maintenance of genetic variation. Genetics. 2013;194:235-44.

54. Meng L, Chen G, Li Z, Yang Y, Wang Z, Wang L. Refugial isolation and range expansions drive the genetic structure of *Oxyria sinensis* (Polygonaceae) in the Himalaya-Hengduan Mountains. Sci Rep. 2015;5:10396.
55. Wu CI. The genic view of the process of speciation. J Evol Biol. 2001;14:851-65.
56. Nosil P, Harmon LJ, Seehausen O. Ecological explanations for (incomplete) speciation. Trends Ecol Evol. 2009;24:145-56.
57. Huang BH, Huang CW, Huang CL, Liao PC. Continuation of the genetic divergence of ecological speciation by spatial environmental heterogeneity in island endemic plants. Sci Rep. 2017;7:5465.
58. Martin NH, Willis JH. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric mimulus species. Evolution. 2007;61:68-82.
59. Anacker BL, Strauss SY. The geography and ecology of plant speciation: range overlap and niche divergence in sister species. Proc R Soc B-Biol Sci. 2014;281:20132980.
60. Zheng H, Powell CM, Rea DK, Wang J, Wang P. Late Miocene and mid-Pliocene enhancement of the East Asian monsoon as viewed from the land and sea. Global Planet Change. 2004;41:147-55.
61. Yao YF, Bruch AA, Mosbrugger V, Li CS. Quantitative reconstruction of Miocene climate patterns and evolution in Southern China based on plant fossils. Paleogeogr Paleoclimatol Paleoecol. 2011;304:291-307.
62. Zachos J, Pagani M, Sloan L, Thomas E, Billups K. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science. 2001;292:686-93.
63. Ferriol M, Picó B, de Córdova PF, Nuez F. Molecular diversity of a germplasm collection of Squash (*Cucurbita moschata*) determined by SRAP and AFLP markers. Crop Sci. 2004;44:653-64.
64. Budak H, Shearman RC, Parmaksiz I, Gaussoin RE, Riordan TP, Dweikat I. Molecular characterization of *Buffalograss* germplasm using sequence-related amplified polymorphism markers. Theor Appl Genet. 2004;108:328-34.
65. Yang J, Vázquez L, Feng L, Liu Z, Zhao G. Climatic and soil factors shape the demographical history and genetic diversity of a deciduous oak (*Quercus liaotungensis*) in Northern China. Front Plant Sci. 2018; 9:1534.
66. Sun Z, Wang Z, Tu J, Zhang J, Yu F, McVetty PBE, Li G. An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. Theor Appl Genet. 2007;114:1305-17.
67. Gutiérrez-Ortega JS, Salinas-Rodríguez MM, Martínez JF, Molina-Freaner F, Pérez-Farrera MA, Vovides AP, Matsuki Y, Suyama Y, Ohsawa TA, Watano Y et al. The phylogeography of the cycad genus *Dioon* (Zamiaceae) clarifies its Cenozoic expansion and diversification in the Mexican transition zone. Ann Bot 2017;121:535-48.
68. Yang XL, Xu QH, Zhao HP, Liang WD, Sun LM. Vegetation changes of the Taihang Mountains since the last glacial. Chin Geogr Sci. 2000;10:261-9.
69. Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R: Common factors drive adaptive genetic variation at different spatial scales in Arabis alpina. Mol Ecol. 2010;19:3824-35.

70. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 1980;8:4321-6.
71. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451-2.
72. Pons O, Petit RJ. Measwring and testing genetic differentiation with ordered versus unordered alleles. Genetics. 1996;144:1237-45.
73. Yeh FC, RC Y, Boyle T. POPGENE Version 1.31. Microsoft windows-based freeware for population genetic analysis. University of Alberta and Centre for International Forestry Research. 1998:11-23.
74. Excoffier L, Lischer HEL: Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10:564-7.
75. Peakall R, Smouse PE. GenALEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics. 2012;28:2537-9.
76. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984;38:1358-70.
77. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870-4.
78. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 2009;9:1322-32.
79. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol. 2005;14:2611-20.
80. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012;4:359-61.
81. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007;23:1801-6.
82. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes. 2004;4:137-8.
83. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008;24:1403-5.
84. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010;11:94.
85. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16:37-48.
86. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29:1969-73.
87. Posada D. jModelTest: Phylogenetic model averaging. Mol Biol Evol. 2008;25(7):1253-1256.

88. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585-95.
89. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147:915-25.
90. Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J-M, Estoup A. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 2014;30:1187-9.
91. Beerli P. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics. 2005;22:341-5.
92. Wilson GA, Rannala B. Bayesian inference of recent migration rates using multilocus genotypes. Genetics. 2003;163:1177-91.
93. Meirmans PG. Nonconvergence in Bayesian estimation of migration rates. Mol Ecol Resour. 2014;14:726-33.
94. Grinnell J. The niche-relationships of the *California Thrasher*. Auk. 1917;34:427-33.
95. Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG. Where is positional uncertainty a problem for species distribution modelling? Ecography. 2014;37:191-203.
96. Borcard D, Gillet F, Legendre P. Unconstrained ordination, numerical ecology with R. Springer. 2011: 115-51.
97. Rosenberg MS, Anderson CD. PASSaGE: Pattern analysis, spatial statistics and geographic exegesis. Version 2. Methods Ecol Evol. 2011;2:229-32.
98. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14:927-30.
99. Wang IJ. Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution. 2013;67:3403-11.
100. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20:289-90.
101. Legendre P, Oksanen J, ter Braak CJF. Testing the significance of canonical axes in redundancy analysis. Methods Ecol Evol. 2011;2:269-77.
102. Vrieze SI. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods. 2012;17:228-43.
103. Blanchet FG, Legendre P, Borcard D. Forward selection of explanatory variables. Ecology. 2008;89:2623-32.
104. Dray S, Legendre P, Blanchet G. packfor: Forward Selection with permutation (Canoco p. 46). R package version 00-7/r58.

Tables

Table 1 The estimated diversity indexes of *Opisthopappus*

Number hapelotypes	Haplotype diversity (*H*d)	Nucleotide diversity (π)	N_a	N_e	*H*	*I*	*PPL*
H1(1)H2(2)H3(1)H4(1)	0.9	0.00083	1.2000	1.1031	0.0631	0.0971	20.00%
H5(1)H6(1)H7(1)H8(1)H9(1)	1	0.00198	1.2923	1.145	0.0912	0.1413	29.23%
H10(1)H11(1)H12(1)H13(1)H14(1)	1	0.00208	1.2000	1.146	0.0803	0.1167	20.00%
H8(1)H15(1)H16(1)H17(1)H18(1)	1	0.00135	1.2308	1.1051	0.0671	0.1056	23.08%
H19(1)H20(1)H21(1)H22(1)H23(1)	1	0.00271	1.1538	1.052	0.0362	0.0603	15.38%
H24(1)H25(1)H26(1)H27(1)H28(1)	1	0.00344	1.3231	1.1861	0.1118	0.169	32.31%
H29(1)H30(1)H31(1)H32(1)H33(1)	1	0.00229	1.2308	1.1345	0.0798	0.1204	23.08%
H5(1)H34(1)H35(1)H36(1)H37(1)	1	0.00271	1.2923	1.156	0.0962	0.1473	29.23%
H22(1)H38(1)H39(1)H40(1)H41(1)	1	0.00146	1.2769	1.1354	0.0843	0.131	27.69%
H2(1)H3(1)H42(1)H43(1)H44(1)	1	0.00292	1.3385	1.1701	0.1063	0.1643	33.85%
H5(1)H30(1)H45(1)H46(1)H47(1)	1	0.00328	1.2615	1.1663	0.0959	0.1425	26.15%
47	0.99327	0.00308	1.7692	1.2532	0.1695	0.2777	76.92%
H48(2)	0	0	1	1	0	0	0.00%
H49(1)H50(2)H51(1)H52(1)	0.9	0.00167	1.2308	1.1234	0.074	0.1134	23.08%
H52(1)H53(3)H54(1)	0.7	0.00115	1.2154	1.116	0.0709	0.1084	21.54%
H50(1)H52(1)H54(1)H55(1)H56(1)	1	0.00094	1.1692	1.1172	0.0662	0.0971	16.92%
H56(1)H57(2)H58(1)H59(1)	0.9	0.00172	1.2462	1.1249	0.0776	0.1197	24.62%
H53(1)H56(1)H60(1)H61(1)H62(1)	1	0.00146	1.2000	1.1277	0.0729	0.1084	20.00%
H56(1)H63(2)H64(1)H65(1)	0.9	0.00115	1.3231	1.1926	0.1116	0.1676	32.31%
H60(3)H66(1)H67(1)	0.7	0.00083	1.2154	1.1292	0.0765	0.1148	21.54%
H50(1)H68(1)H69(2)H70(1)	0.9	0.00094	1.2769	1.1457	0.0879	0.1348	27.69%
H50(1)H52(1)H55(3)	0.7	0.00042	1.3385	1.1617	0.0999	0.1561	33.85%
H50(2)H52(1)H53(1)H71(1)	0.9	0.00479	1.2000	1.0967	0.0601	0.0937	20.00%
H50(1)H52(1)H56(1)H72(2)	0.9	0.00146	1.1692	1.0946	0.0557	0.0847	16.92%
H51(1)H53(1)H73(1)H74(1)H75(1)	1	0.00198	1.1077	1.0395	0.0271	0.0443	10.77%
28	0.95399	0.00178	1.7385	1.246	0.1592	0.2585	73.85%
ih 75	0.98571	0.0073	1.9385	1.3985	0.2460	0.3887	93.85%

Na = Observed number of alleles, *N*e = Effective number of alleles, *H* = Nei's gene diversity,

I = Shannon's Information index, *PPL* = The percentage of polymorphic sites.

Table 2 Analysis of molecular variance (AMOVA) based on pairwise differences for *Opisthopappus*

SNP	Source	df	SS	MS	Est. Var.	%	Fixation Indices
	Among species	1	548.9707	548.9707	9.3417	80%	F_{CT}=0.8003
	Among populations within species	22	96.6499	4.3932	0.5331	5%	F_{SC}=0.2287
	Within populations	93	167.2000	1.7978	1.7978	15%	F_{ST}=0.8460 (P<0.001)
	Total	116	812.8205		11.6727	100%	
InDel	Source	df	SS	MS	Est. Var.	%	Fixation Indices
	Among species	1	321.7912	321.7912	5.2026	44%	F_{CT}=0.4364
	Among populations within species	22	402.0003	18.2727	2.9867	25%	F_{SC}=0.4444
	Within populations	93	347.2000	3.7333	3.7333	31%	F_{ST}=0.6869 (P<0.001)
	Total	116	1070.9915		11.9226	100%	

F_{CT}, genetic differentiation among groups

F_{SC}, genetic differentiation among populations within groups

F_{ST}, genetic differentiation among populations

Table 3 Summary of the Mantel test and multiple matrix regression with randomization (MMRR) between the genetic (gen), geographic (geo), and environmental (env) distances

	Mantel test		MMRR		
	r	P	coefficient	r^2	P
geo vs env	0.0161	0.421	0.0702	0.0003	0.882
gen vs geo	0.5039	0.001	0.0029	0.2539	0.001
gen vs env	0.1512	0.024	0.0038	0.0228	0.016
gen vs env (conditioned on geo)[a]	0.1656	0.011	--	--	--
gen vs (geo*env)[b]	--	--	geo:0.0029	0.2744	0.001
			env:0.0036		0.019

a: The partial Mantel test.

b: The joint effect of both environmental and geographic distances in MMRR.

Table 4 Summary of partial dbRDA, showing the significance of climatic variables (constrained factors) for explaining the variation in the genetic components

					GLM for the distr. of climatic variables along ordination axes						
	Inertia (Var)[a]	Proportion	adj R^2	P	t (axis 1)	Pr (>|t|)	t (axis 2)	Pr (>|t|)	Adj R^2	F	P
tioned tude + tude)	0.3179	11.11%	0.0985	0.04							
rained	1.7635	61.64%	0.5099	0.01							
	0.2238	7.82%	0.0503	0.08	2.1220	0.0454	-0.6110	0.5480	0.1178	4.0700	0.0560
	0.7162	25.03%	0.2374	0.02	-4.8750	0.0001	0.1420	0.8880	0.4768	21.9600	0.0001
	0.1473	5.15%	0.0335	0.08	0.2100	0.8360	1.4050	0.1740	-0.0407	0.1002	0.7546
	0.3411	11.92%	0.0958	0.02	-2.4420	0.0231	-2.2140	0.0375	0.2047	6.9200	0.0153
	0.3355	11.72%	0.0930	0.06	2.3250	0.0297	-0.0510	0.9600	0.1535	5.1700	0.0331
nstrained	0.7794	27.24%									
	2.8607	100%									

a: Inertia is the mean squared Euclidean distance

Table 5 Importance of environmental variables explaining genetic variation

'ordistep' function in R package 'vegan'					
		Df	AIC	F	P
step 1	bio6	1	13.628	18.3818	0.005
	bio3	1	24.14	4.0602	0.035
	bio16	1	23.609	4.6432	0.045
	bio14	1	23.393	4.8833	0.050
	bio8	1	28.088	0.1071	0.885
		Df	AIC	F	P
step 2	bio14	1	9.6132	5.9817	0.020
	bio8	1	10.712	4.7742	0.020
	bio16	1	14.6765	0.8497	0.380
	bio3	1	14.7414	0.7906	0.450
		Df	AIC	F	P
step 3	bio8	1	10.866	0.6324	0.445
	bio16	1	11.031	0.4909	0.520
	bio3	1	11.29	0.2711	0.795

'forward.sel' function in R package 'packfor'[a]					
	R^2	R^2 Cum	Adj R^2 Cum	F	P
bio6	0.4552	0.4552	0.4304	18.3818	0.001
bio14	0.1208	0.5760	0.5356	5.9817	0.014

a: Procedure stopped (adjR^2 thresh criteria) adjR^2cum = 0.535595 with 2 variables (superior to 0.509899)

Figures

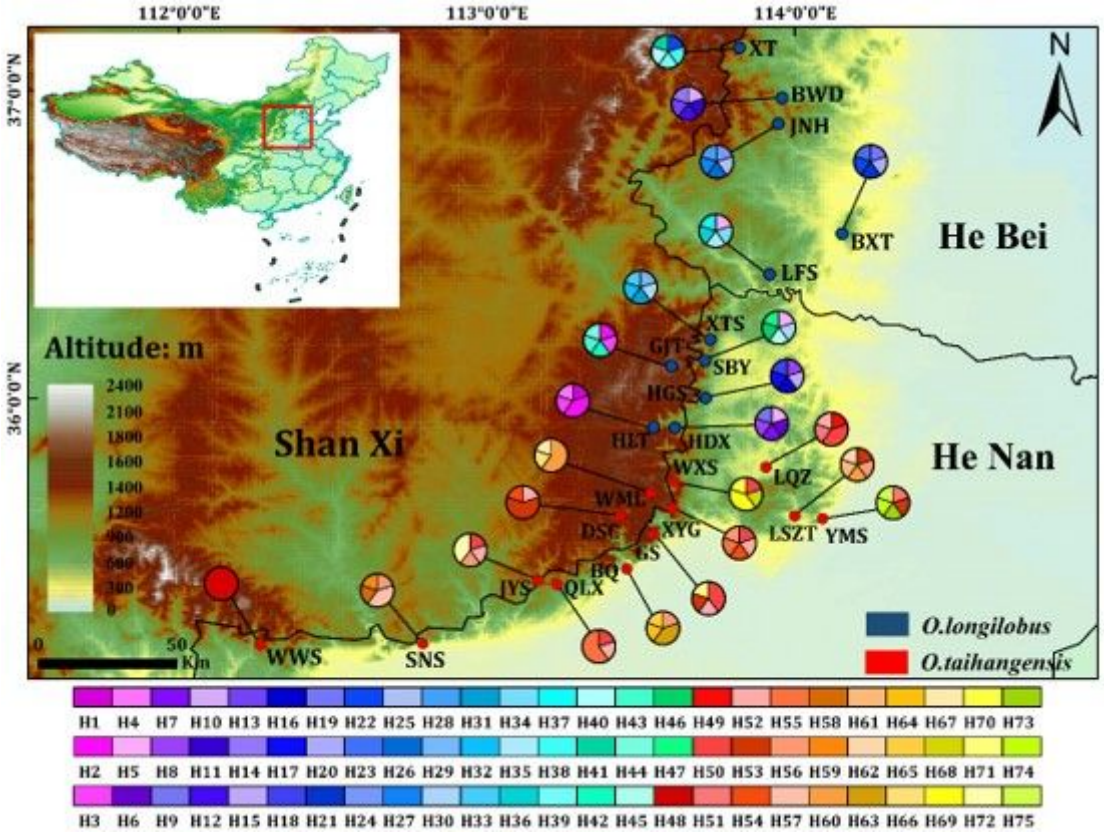

Figure 1

The sampling site of Opisthopappus populations and distribution of haplotypes. Red circles mean O. taihangensis populations and blue circles mean O. longilobus populations. For population abbreviations, see additional file 10 for details.

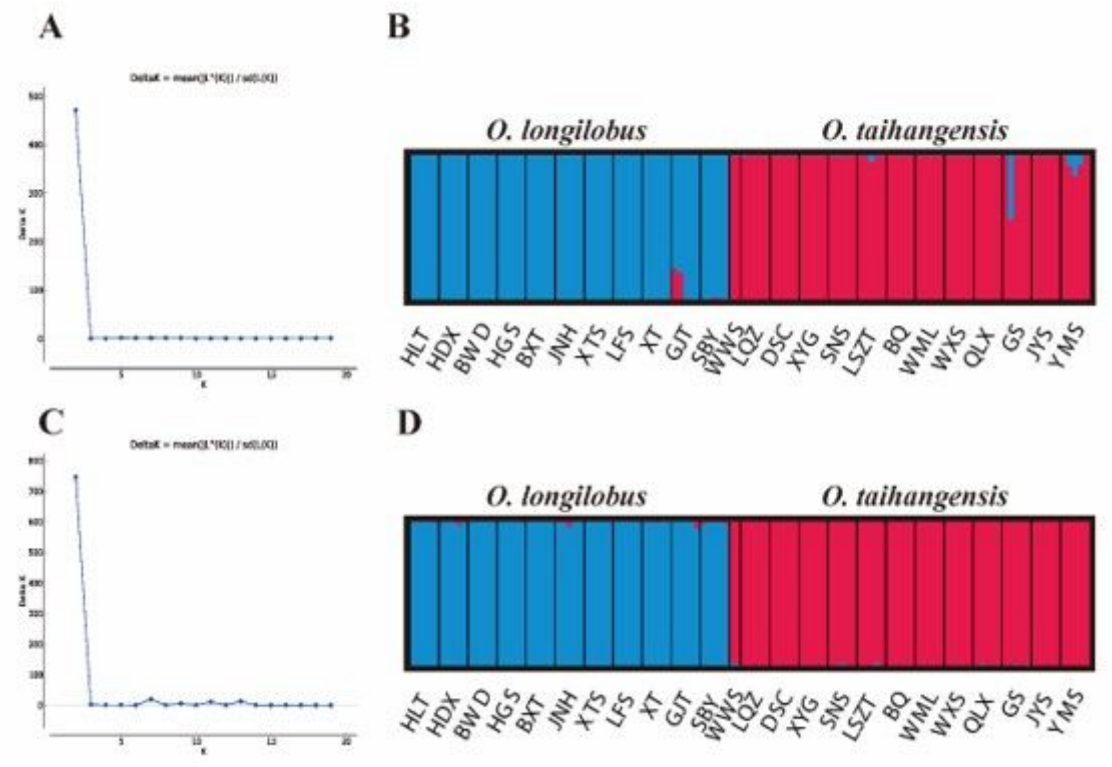

Figure 2

Results of the Bayesian clustering analysis conducted using STRUCTURE. The ΔK plot conducted by Structure Harvester showed that K = 2 got the highest ΔK value. (A): ΔK plot of SNP. (B): Estimated genetic structure for K = 2 based on SNP. (C): ΔK plot of InDel. (D): Estimated genetic structure for K = 2 based on InDel.

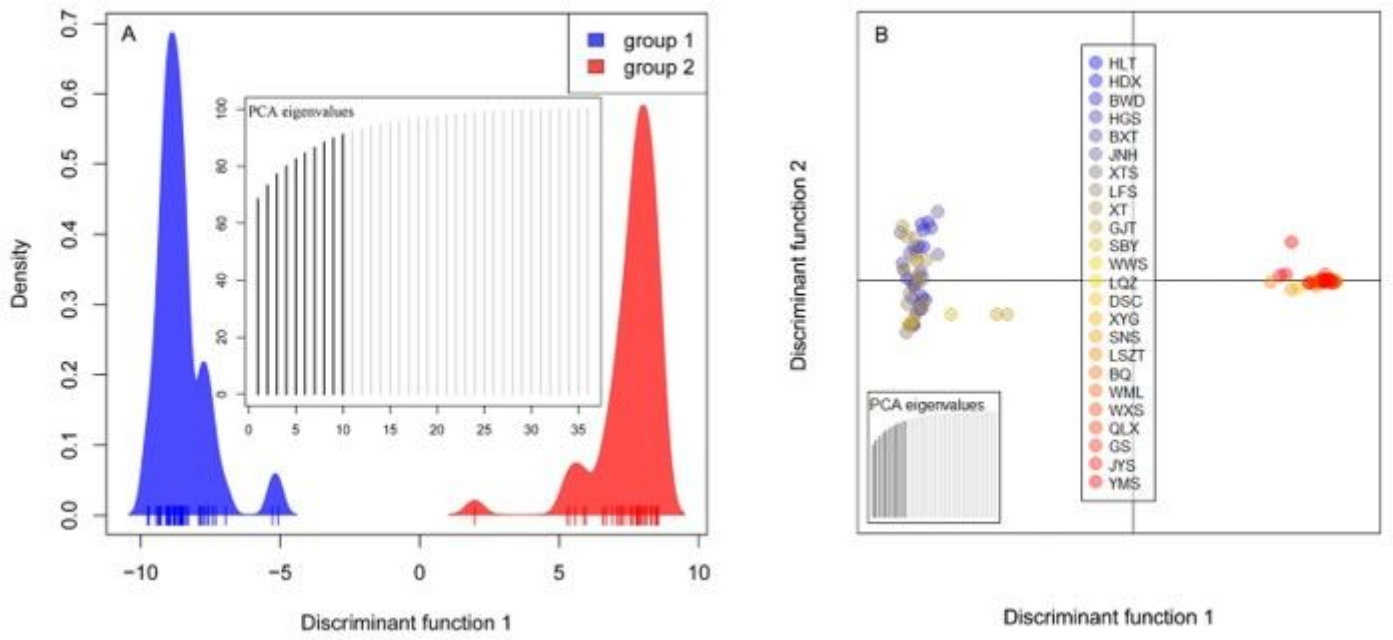

Figure 3

Population structure revealed by the discriminant analysis of principal components (DAPC). (A): The DAPC for separating the species O. longilobus (blue) and O. taihangensis (red). (B): The DAPC for separating populations of O. longilobus and O. taihangensis. The blue series belongs to O. longilobus and the red-color series are O. taihangensis.

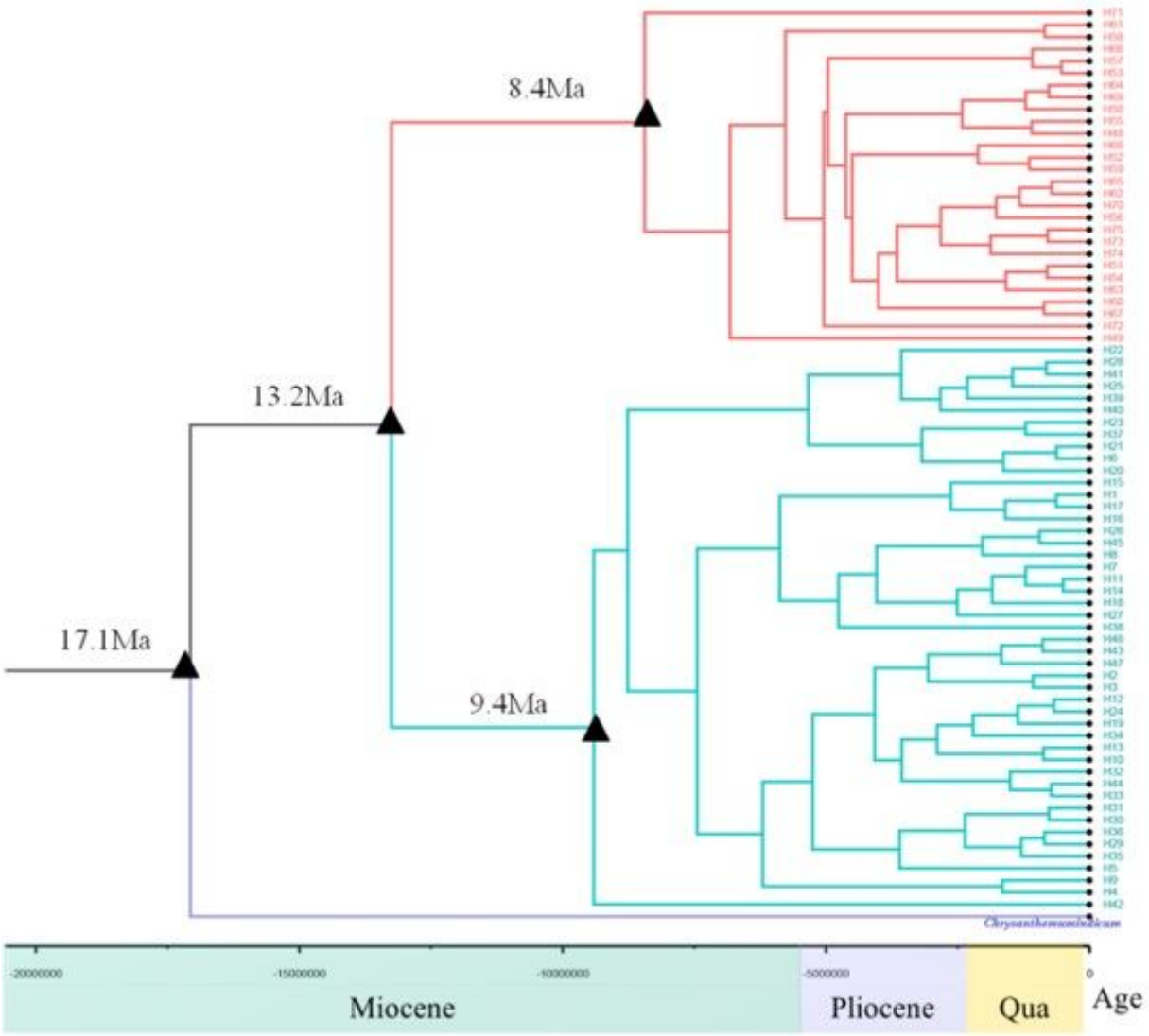

Figure 4

BEAST-derived chronogram of Opisthopappus based on haplotypes. The red branches represent haplotypes of O. taihangensis and the blue branches represent haplotypes of O. longilobus. The divergence time between the Opisthopappus and the outgroup (Chrysanthemum indicum) was estimated to 17.1 Ma, in early Miocene. The divergence time between O. taihangensis and O. longilobus was estimated to 13.2 Ma, in middle Miocene. The initial haplotype divergences time of O. longilobus and O. taihangensis were estimated to 8.4 Ma and 9.4 Ma, respectively. The diversification time of specific haplotypes within species was mainly in Quaternary.

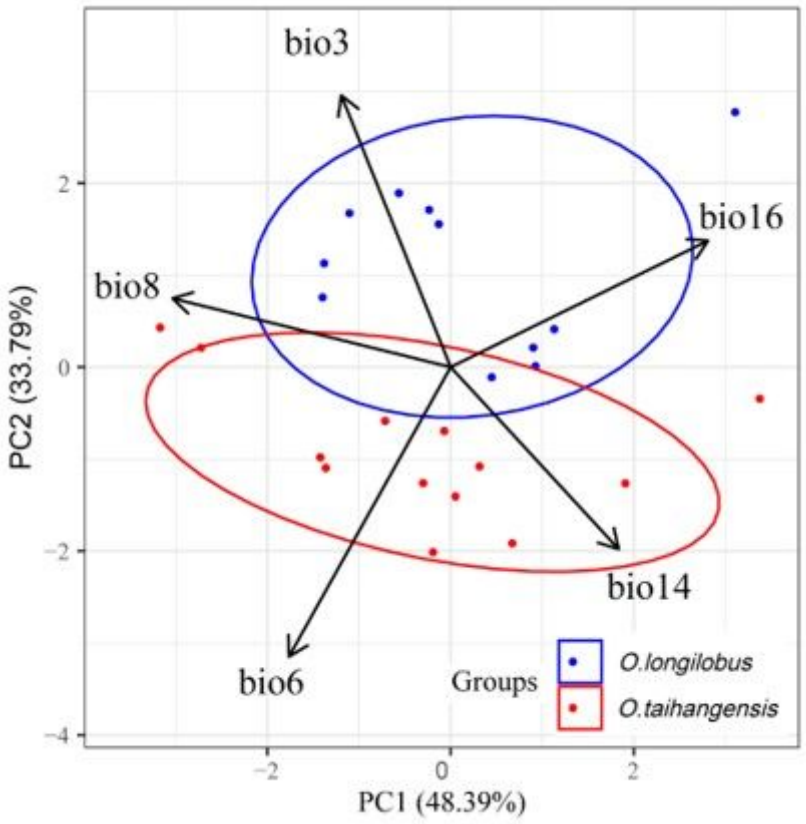

Figure 5

Principal component analysis of the climate factors of O. longilobus and O. taihangensis.

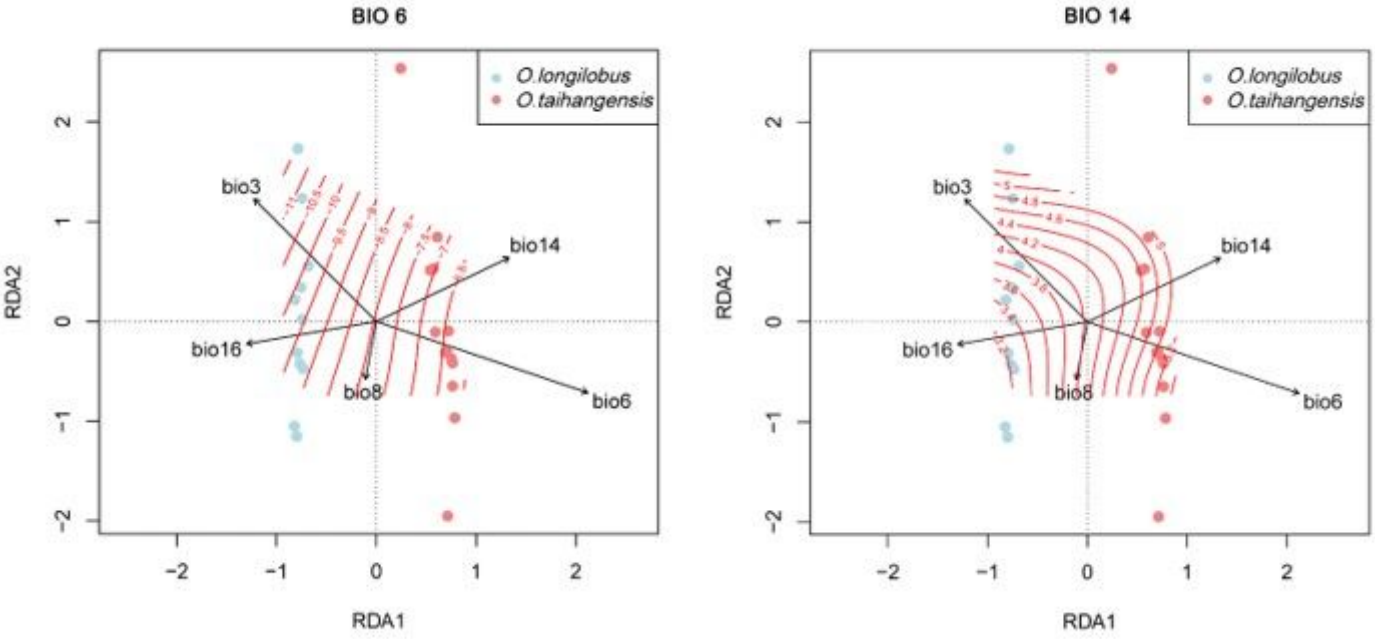

Figure 6

Scatter and ordisurf plots of the dbRDA for the two important explanatory climate variables (bio6 and bio14).

Supplementary Files

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

- Additionalfile9FigureS6.docx
- Additionalfile2FigureS2.docx
- Additionalfile5TableS2.docx
- Additionalfile8FigureS5.docx
- Additionalfile4TableS1.docx
- Additionalfile6FigureS4.docx
- Additionalfile3FigureS3.docx
- Additionalfile1FigureS1.docx
- Additionalfile11TableS5.docx
- Additionalfile10TableS4.docx
- Additionalfile7TableS3.docx