Cytotype Distribution and Chloroplast Phylogeography of *Actinidia Chinensis* Complex

Zhi Wang
Wuhan Botanical Garden
Caihong Zhong
Wuhan Botanical Garden
Dawei Li
Wuhan Botanical Garden
Chunlin Yan
Wuhan Botanical Garden
Xiaohong Yao (yaox@wbgcas.cn)
Wuhan Botanical Garden
Zuozhou Li
Wuhan Botanical Garden

Research article

Keywords: phylogeography, polyploidy, genetic diversity, ecological niche models, subtropical China, Actinidia chinensis complex

Posted Date: October 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-64406/v1

Version of Record: A version of this preprint was published at BMC Plant Biology on July 6th, 2021. See the published version at https://doi.org/10.1186/s12870-021-03099-y.

Abstract

Background: Previous plant phylogeography in subtropical China mostly focused on rare and endangered species, whereas few studies conducted on the taxa with relatively wide distribution, especially for polyploid species. We used *Actinidia chinensis* complex, a widespread geographically woody liana of variable ploidy in subtropical China, to investigate its cytotype and haplotype distribution pattern based on three chloroplast fragments (*ndh*F-*rpl*132, *ps16*-*trn*Q and *trn*E-*trn*T). Macroevolutionary, microevolutionary and niche modeling tools were also combined to disentangle the origin and the demographic history of the species or cytotypes.

Results: The ploidy levels of 3569 individuals from 128 populations sampled throughout the species distribution range were estimated using flow cytometry. The widespread cytotypes were diploids followed by tetraploids and hexaploids, while triploids and octaploids occasionally occurred in a few populations. Thirty-one chloroplast haplotypes were detected and the genetic diversity and genetic structure were great differences between the var. *chinensis* and var. *deliciosa*. Our results also revealed these two varieties (or ploidy races) inhabit significantly different climatic niche spaces. Ecological niche models (ENMs) indicate that all varieties' ranges contracted during the Last Interglacial (LIG), and expanded eastward or northward during the Last Glacial Maximum (LGM).

Conclusions: Pliocene and Plio-Pleistocene climatic fluctuations and vicariance may have played key roles in shaping the current population structure and historical demography in the *A. chinensis* complex. The polyploidization process also played an important role in the historical demography of the *A. chinensis* complex through improving their adaptability to environmental changes.

Background

In the past decades, many phylogeography studies have focused on the Qinghai-Tibetan Plateau and adjacent areas in China [1–4]. Moreover, as an area harboring many ancient endemic genera and families, the subtropical region in China has recently attracted more attention due to the complex topography and fluctuated paleoclimate of the area in recent years [5–7]. Nevertheless, most of these phylogeographic studies in these two hotspots of China have mainly focused on rare and endangered species [8–11], with few cases examined in the species with a relatively broad range [12, 13], especially for polyploid species. Polyploidization is a common phenomenon in higher plants and plays an important role in plant diversification and adaptation to novel environments [14]. Although the wide recognition of the prevalence of polyploidy and there was also much more cytological evidence for recent whole-genome duplication(s) in many plant groups [15, 16], the studies on the phylogeography of polyploid species in these two hotspots in China [17–19] were rarer than those in other hotspots in the world (e.g. North America and Alps in Central Europe).

Actinidia chinensis complex is a widespread geographically woody liana of variable ploidy in subtropical China and also planted widely around the world as an important horticultural fruit tree crops. In the newly

revised infrageneric system of *Actinidia*, the species complex *Actinidia chinensis* is composed of three varieties, *A. chinensis* var. *chinensis* (hereafter: var. *chinensis*), *A. chinensis* var. *setosa*, and *A. chinensis* var. *deliciosa* (hereafter: var. *deliciosa*) [20]. Previous studies have shown that the vast majority of var. *chinensis* and var. *deliciosa* can be distinguished by morphological characteristics such as the presence of overwintering buds or the hair types covering on flowering branchlets and fruits [21]. However, there are still a certain number of morphological intermediates of these two varieties that are identified in their distribution overlap region in central China [22–24]. In terms of distribution, var. *chinensis* grows mainly in eastern and central China, while var. *deliciosa* is more inland in central and western China [25]. Although, they coexist usually in many localities of the overlap region but slightly separated vertically with var. *chinensis* being found at lower altitudes. The previous cytogeographic analysis has revealed that ploidy types of *A. chinensis* complex vary from diploid to hexaploid and these two varieties were also composed by different cytotype (2n = 2x = 58, var. *chinensis*: 2x and 4x, var. *deliciosa*: 6x with few 4x) [24]. The existing population genetic studies of the *A. chinensis* complex were based on relatively fewer sample sets and did not cover the whole distribution range of this species [21, 23, 24, 26]. Thus, the genetic diversity and population divergence history of the *A. chinensis* complex with different ploidies are still unclear.

In addition, the paleoecological reconstruction of forest biomes can supply fundamental guidance for testable phylogeographic hypotheses, while they cannot provide details of population history [1, 27]. Ecological niche models (ENMs), which source authorizes hoary arise distributions footing be worn to augment the limited fossil record in East Asia [28]. Combined with molecular information, ENMs can strengthen our familiarity with the material dimension of population dynamics [29, 30].

In this study, the flow cytometric measurement (FCM) [31] was used to identify cell types of the *A. chinensis* complex, and the cpDNA markers were used to investigate the genetic diversity and population structure of the *A. chinensis* complex. Besides, the mismatch distribution analysis (MDA) and the ecological niche models (ENMs) were used to investigate the evolutionary history and demographic structure of their natural populations in subtropical China. The main objectives of our study were to (a) reveal the cytotype distribution pattern; (b) evaluate the level of genetic diversity and the degree of genetic differentiation among populations for each variety; (c) speculate on the origin of polyploidy and explore the demographic history of the *A. chinensis* complex. To summarize, our goal is to explore patterns and mechanisms of population diversity and historical demography of a typical polyploidy plant in the subtropical forest that generally corresponding to the Metasequoia Flora (Sino-Japanese Kingdom), one of the two major floristic regions in the East Asian Floras [7].

Results

Cytotype diversity and distribution

Five ploidy types (di-, tri-, tetra-, hexa- and octaploid, Additional file 1: Figure S1) were identified (Additional file 2: Table S1). All the diploids belonged to var. *chinensis*, whereas all the hexaploids

belonged to var. *deliciosa*. The tetraploids were shared in both var. *chinensis* and var. *deliciosa* pool. Sixteen triploids having similar morphological characteristics with var. *chinensis* were observed in the diploid, tetraploid, and mixed populations (2x, 4x), while rare octaploids were found in two hexaploid var. *deliciosa* populations, respectively.

The cytotype distribution was shown in Figure 1. Diploid var*. chinensis* and hexaploid var*.deliciosa* exhibited clear spatial segregation along with east China to west China, and their main distribution overlaps in mountainous regions of the second step of China's terrain. The tetraploids distribution area was mainly in Wuyi mountain and Mufu mountain in east China and Xuefeng mountain in the Midwest China. The subsequent phylogenetic analysis also showed that two parts of the tetraploids had belonged to two different lineages (Figure 3a) which have differentiated at 2.22 Ma.

Genetic diversity and population structure

Three cpDNA fragments (*ndh*F-*rpl*132, *ps*16-*trn*Q, and *trn*E-*trn*T) across 608 sampled individuals from 49 populations were successfully sequenced. The total length of the concatenated sequence was 1854 bp. A total of 35 polymorphic sites and 31 haplotypes were identified (Additional file 2: Table S2). Twenty-eight haplotypes were present in var*.chinensis* and three haplotypes were present in var*.deliciosa*. Four haplotypes were shared by var*.* chinensis and var*. deliciosa* (Figure 2, Additional file 2: Table S3). The total genetic diversity of var*. chinensis* (H_T = 0.920, H_S = 0.470) was higher than that of var. *deliciosa* (H_T = 0.552, H_S = 0.380). The coefficient of genetic differentiation N_{ST} was significantly larger than G_{ST} for var*. chinensis*, while the results of var*.deliciosa* were exactly opposite (Table 1).

None spatial population structure was detected by the SAMOVA analysis of cpDNA haplotype from 49 populations, due to the F_{CT} values did not have significant peaks as the value of *K* increased from 2-20 (Additional file 1: Figure S2). Therefore, the sampled populations were divided into two groups by taxonomy for the AMOVA analysis. The AMOVA analysis (Additional file 2: Table S6) related that 25.06% of the molecular variation was distributed between varieties (F_{CT} = 0.251, P < 0.01). The inter-variant population fixation indexes (F_{ST}) were 0.654 and 0.352 (P < 0.01) for var*.chinensis* and var*.deliciosa*, respectively. The isolation-by-distance (IBD) tests based on the pairwise $F_{ST}/(1- F_{ST})$ genetic distances against the geographic distances detected no significant correlation for the cpDNA (R^2 = 0.1234, P = 0.988) among the 49 populations of the *A. chinensis* complex.

Phylogenetic relationships and Divergence dating

Phylogenetic analysis of the ingroup concatenated cpDNA haplotype data sets recovered three haplotype lineages (lineage 1, 2, and 3, Figure 3a), which was highly consistent with the phylogenetic network based on the 31 haplotypes (Figure 3b). The haplotypes lineages also revealed the district distribution range (Figure 2). The haplotypes in lineage 1 were mainly found in mountainous regions of the second step of China's terrain where the two varieties overlapped. The haplotypes that belong to var*. chinensis* (2x), except H16 were shared by var*.chinensis* and var*. deliciosa*. The haplotypes in lineage 3 which also

belong to var. *chinensis* (2x) were found in eastern areas, such as Wuyi mountain, Dabie mountain, and Luoxiao mountain. Finally, almost all polyploid haplotypes were found in lineage 2. The dominant haplotype H10 and H11 occurred in 154 and 111 individuals, respectively. Although these two haplotypes were shared by var.*chinensis* and var.*deliciosa*, they still mainly appertain to var. *deliciosa*. Half haplotypes (16/31) were restricted to a single population. For example, in lineage 3, H3, H6, and H8 were found exclusively in Dabie mountain and Huang mountain, while H28, H30, and H31 appeared only in the Zhejiang-Fujian Hilly Region.

The time-calibrated evolutionary tree (Figure 3a) showed that the divergence time between the *A.chinensis* complex and the outgroup (*A. eriantha*) could date back to 11 million years ago (Ma) during the late Miocene. For the 31 cpDNA haplotypes of *A.chinensis* complex, the mean crown-group age of the *A.chinensis* complex was ca. 4.43 Ma (95% HPD 2.81–6.39 Ma, node A), thereby indicating the recent diversification of the cpDNA haplotypes of the *A.chinensis* complex during the Pliocene. Two polyploid haplotype lineages were diverged at ca. 2.22 Ma (95% HPD 1.14–3.57 Ma, node C), which suggests the var. *deliciosa* was evolved into a single lineage.

Demographic history and ancestral geographical area

The mismatch distributions of group 2 and 3 were unimodal (Additional file 1: Figure S3), and the *F*s test and Tajima's *D* test of group 2 and 3 were also negative, although only group 2 were highly significant ($P < 0.05$, Table 2), closely fitted to the expected distribution under the sudden expansion model. For the other two groups (group 1 and 4), there were no explicit signals of population expansion or equilibrium, as evidenced from the neutral tests and mismatch analysis (Table 2; Additional file 1: Figure S3), although the variance of the Sum of square deviations and Harpending's raggedness index tests (*SSD* and *HRI*, $P > 0.05$; Table 2) failed to reject the spatial expansion model. The tua values (τ) presented a rough estimate when the population expansion started. The approximate expansion times (year) for groups 2 and 3 were measured as 0.084 (0.033-0.136) Ma BP, 0.137 (0.081-0.162) Ma BP, respectively.

The BBM (Bayesian Binary MCMC) analysis of ancestral distribution areas (Figure 4) revealed that two vicariance and twelve dispersal events were discovered respectively. Based on the topology of the intraspecific chronogram, a likely ancient vicariance event (node □, EA, and CE) of two sister lineages was founded. Besides, a subsequent vicariance event (node □) between populations of tetraploid var. *chinensis* in eastern (EA) and tetraploid/hexaploid var. *deliciosa* in central-western (CE and WE) China was also presented. These events were then followed by dispersal main from the east (EA) to the central (CE), the central (CE) to the west (WE) or the south (SO). Recent colonization events from the east (EA) to the central (CE) were inferred, based on the genetic admixture (H1, H2, and H7) in the adjacent mountains (e.g. Qinling and Dabie Mountains) or between the distant mountains (e.g. Qinling and Wuyi Mountains), with multiple expansions at different times (Fig. 4). For the diploid varieties, the mountains of eastern China (EA) were inferred as the most likely ancestral area. The ancient haplotypes (H12, Figure 3b) identified in haplotype network analysis were also distributed in the population ZJTS and FJGT of eastern China. However, for the polyploid varieties, the mountains of central China (CE) such as Xuefeng

mountain were deduced as the most likely ancestral area, which was also supported by the results of ENM analysis. No extinction event was discovered in any lineage by BBM analysis.

Ecological niche modeling and niche comparation

The AUC (the Area Under the Curve) values for each variety and climate scenario were high (AUC= 0.978/0.982), indicating that all models performed well in predicting the suitable habitat for each variety. The current distributional predictions were good representations of the actual distribution of both varieties. The main geographic distributional areas for both varieties were predicted within the subtropical region in China, mainly in the mountains. The elevation of distribution regions for var*. deliciosa* was often higher than that of var. *chinensis*. The distribution areas of two varieties overlapped in mountainous regions of the second step of China's terrain. The predicted distribution range for var.*deliciosa* during the LIG was more southward than that of the present (Figure 5 a, e). Var. *chinensis* also had a more restricted distribution range during the LIG than that in the present, and the range was more broken than the current distribution (Figure 5 b, f). The predicted species distribution during the LGM for var. *chinensis* and var. *deliciosa* was both significantly larger than the predicted distributions under the LIG or current conditions (Figure 5), suggest the distribution ranges of the two varieties had expanded during the LGM.

Observed measures of niche-overlap and identity (D=0.433 and I=0.708, see Figure 6a) were lower than null distributions for var. *chinensis* vs. var. *deliciosa*, suggesting high niche differentiation between these two varieties. In addition, more detailed discrepancies can be found in the niches of three different cytotypes (Additional file 1: Figure S4). Thus, each variety or cytotype occupied a different niche with distinct environments, which was also supported by the results of PCA analysis based on the data from elevation and 19 bioclimatic variables (Figure 6b).

Discussion

High ploidy diversity in *A. chinensis* complex

To our knowledge, this study represents the most comprehensive investigation of the ploidy variation of *A. chinensis* complex by examining nearly 130 populations and multiple individuals per population (over 30 in total) across their whole distributed regions. Var. *deliciosa* predominantly comprises tetraploids and hexaploids, whereas var. *chinensis* consists of diploids and tetraploids, which are largely consistent with the results of Li et al [24] where they only studied the populations sampling from Hunan and Guizhou provinces, China. For *A. chinensis* complex, hexaploids centered in the western and diploids and tetraploid in the eastern portions of their ranges. For var. *deliciosa*, the most frequent and widespread cytotype is hexaploid, followed by tetraploid. For var. *chinensis*, diploid is geographically most widespread followed by tetraploid. Compared to the previous study [24], octaploid var. *deliciosa* is observed in the present study. In addition, natural triploid var. *chinensis* is described for the first time. The low frequencies occurrence of tetraploids within diploid-tetraploid populations suggest the probability of interaction between diploids and tetraploids in the contact zone and highlights the potential for contemporary gene flow from diploid to polyploid populations.

In total, six different ploidy combinations were observed, including two majority cytotypes (2x + 4x, 4x + 6x,) and four minority cytotype (2x + 3x , 2x + 3x + 4x, 2x + 4x + 6x, 2x + 3x + 4x + 6x) (Figure 1). The coexistence of cytotypes at the infraspecific or even intrapopulation levels has been documented for many plant species [32]. The maintenance of multiple cytotypes within the same sites may be explained by the divergence in pollinator spectra and/or visit frequency, different flowering time, and other factors associated with reproductive isolation [24]. For example, our field investigation shows diploids often bloom slightly earlier than tetraploids and hexaploids, which suggests a reproductive barrier between cytotypes in the overlapping niches. The overall cytotype distribution pattern may arise from divergent adaptive abilities to the local environment. Generally, polyploid species have been suggested to be more adaptive to new ecological conditions than the ancestral diploid [33]. The overall geographical pattern of cytotype segregation in the species complex implies kiwifruit species enhanced their adaptability to harsh climate conditions (e.g. colder climate) by reaching a higher ploidy level [24].

Multiple origins and long-distance dispersal of polyploids

Despite autopolyploids being difficult to identify, several lines of evidence, such as isozymes and DNA sequences derived from the polygalacturonase gene, suggest that tetraploid var. *chinensis* is autopolyploid and hexaploid var. *deliciosa* is allohexaploid [34, 35]. Besides, in this study 31 haplotypes were detected in *A.chinensis* complex, of which 19 were only found in diploids, three in tetraploids, three in hexaploids, three shared by diploids and tetraploids, two shared by tetraploids and hexaploids, and one shared by diploids, tetraploids, and hexaploids. If we assume that the polyploids were produced by diploids with the same cpDNA haplotypes, and each haplotype has only evolved once (without parallel evolution), then we can conclude that in the evolutionary history of the tetraploids in the *A. chinensis* complex had experienced at least four independent origins; and the hexaploids had also experienced at least three independent origins, of which tetraploid might act as a bridge. Previous studies have also proposed that another hypothesis of sharing multiple haplotypes between diploids and tetraploids (polyploids) resulted from the hybridization among the different ploidy race [36, 37], where plastids and nuclear genes can be transferred to tetraploids by unreduced gametes or triploid bridges [38-41]. In the present study, a small number of triploid individuals were found, and these triploid individuals mainly existed in the mixed 2x-4x populations. However, according to our current molecular data, it is not enough to distinguish this hypothesis. But in this study, we still admit that the detection of one or more polyploid origins may reflect the transfer of plastids from diploid to polyploid through intraspecific or interspecific hybridization. As a result, our data suggest at least seven independent polyploid origins occurred in the *A. chinensis* complex. Multiple independent origins were also observed in the *Allium przewalskianum* of the Qinghai-Tibet Plateau and adjacent regions [17].

Although many studies have shown that in situ formations of polyploid may occur frequently, for example, in diploid populations, tetraploids and rare triploids can be occasionally observed [36, 37, 42, 43]. However, it should be emphasized that for the tetraploids in the East China populations, we considered that they originated mainly in the southwest mountainous area as same as other polyploids. After origination, they then extended eastward into the mountainous region of East China through long-

distance dispersal (LDD), in which Luoxiao mountain acted as the middle bridge. The evidence is mainly based on the following two aspects: first, haplotype network and phylogenetic analysis show that the tetraploids in the East gather in a clade and also closely related to the main haplotypes in the Midwestern mountainous region; second, based on the BBM results in the tetraploid populations in the east may be vicariance and dispersal from the Midwestern mountainous region. LDD events can be traced back to the late Pliocene 2.22Ma (1.14-3.57), based on the results of the Bayesian dating (Figure 3a). LDD events have been also revealed in *A. eriantha*, a related species of *A. chinensis* (Guo et al., unpublished data).

Genetic divergence between var. *chinensis* and var. *deliciosa*

In this study, 31 haplotypes were obtained by sequencing the cpDNA of 49 populations of *A. chinensis* complex, and the genetic diversity (H_T) at the species level was 0.884. Among them, the chloroplast genetic diversity (H_T) of var. *chinensis* was 0.920, which was significantly higher than that the relict plants in this region, such as *Cercidiphyllum japonicum* (0.757) [44], *Davidia involucrata* (0.882) [45], and *eupleea pleiosperma* (0.893) [46]. However, the genetic diversity (H_T) of var*. deliciosa* was 0.552, which is significantly lower than that of other species distributed in the same region. Although the evolution rate of the chloroplast genome is generally slow and the level of intraspecific genetic diversity is low [47], there is still a large number of genetic variants accumulated in the long evolutionary history. Var. *chinensis* has evolved about 4.43 million years in China, while var*. deliciosa* has relatively less differentiation time just about 2.22 million years (Figure 3). Moreover, var. *chinensis* has a wide distribution area with high environmental heterogeneity, which results in an increased probability of isolation, drift, and mutation [46]. While var*. deliciosa* is also widely distributed, but its suitable habitats are relatively less than that of var. *chinensis* (see below). In addition, previous studies have also found that widespread species tend to have a higher level of genetic diversity than local species [44]. The same distinction also exists in their genetic structure, var*. chinensis* has the obvious genetic structure (G_{ST} / N_{ST} = 0.490/0.591, P < 0.01), while var*. deliciosa* does not have obvious genetic structure (G_{ST} / N_{ST} = 0.312 /0.252, P > 0.01).

Pliocene intraspecific divergence andpopulation dynamic history

This Global climate cooling is hypothesized to be a key trigger for climate awning in East Asia, which enhances the monsoon climate in East Asia simultaneously [48, 49]. The rapid rise on the eastern edge of the Qinghai-Tibet Plateau in the mid-Pliocene (c. 3.6 Ma) led to dramatic changes in the landscape of southwestern China [49, 50]. The estimated divergence time between the two major lineages of the *A. chinensis* complex was 4.43 Ma (95% CI: 2.81-6.39; see node A in Figure 3), which coincides with the accelerated rate of cooling of the early Pliocene of the Tertiary period [10]. Therefore, we speculate that climate and geological changes since the Pliocene may play important roles in the habitat fragmentation and the geographical barrier effect of the formation of gene flows, which in turn led to early spectral differentiation of the *A. chinensis* complex. Such isolation and differentiation caused by geological and climate change were also used to explain similar patterns of east-west differentiation in other plants from subtropical China such as *Primula Obconicain* [51], *Cyclocarya paliurus* [52], and *Quercus acutissima* [53].

The BBM result of cpDNA and the distribution of ancient haplotype (H12) revealed that the most ancient area of var*. chinensis* and eastern refugium is located in Fujian hilly region (Figure 2 and Figure 4). The moist and warmer climate conditions in Fujian hilly region enhance the survival of warm-temperate deciduous forest species since the LGM [1]. Restricted to Fujian hilly region in the LIG period, var*. chinensis* then rapidly migrated eastward, westward, and northward from this refugial area, reaching the vast regions since the LGM. In addition, seed dispersal mechanism (animal dispersal), polyploidization, and interspecific/intraspecific hybridization may also contribute to its rapid spread and expansion [54].

The ENMs analysis based on Paleoclimate shows that the distribution range of the *A.chinensis* complex has the trend of the Last Interglacial contraction and the Last Glacial Maximum expansion, which is mutually supported by the results of population dynamic analysis. This dynamic pattern is the same as that of *Pinus kwangtungensis* [55] and *Emmenopterys henryi* [9], which belong to the region of subtropical China. The *A. chinensis* complex is a mountain deciduous broad-leaved plant distributed in the whole subtropical zone of China with an altitude of 200-2500 m [25]. Previous studies indicated that changes in temperature can cause mountain species to migrate up and down at mountain altitudes [56, 57]. In the LIG period, the temperature was 2°C higher than that of the present [58], the mountain species will move up to higher altitudes, and the distribution area of species will be more fragmented, even disappear because of the limit of mountain altitude. However, in the LGM period, these species will move down and get more suitable habitats with the decrease of temperature, which shows the expansion of the population, and even the reunion of groups at different mountain ends of the lowland. Besides, polyploidization promoted species to obtain a broader niche, which has been verified in previous studies [59-62]. In this study, the ecological niche differentiation between var*. chinensis* and var*. deliciosa* has been confirmed. In fact, our field survey data directly show that the altitude distribution range of var. *deliciosa* (including tetraploid of var*. chinensis*) is higher than that of var*. chinensis* (mainly diploid). Therefore, based on the MDA analysis of demographics, two groups (groups 2 and 3, Additional file 1: Figure S3) including polyploid individuals show a clearer history of population expansion.

Polyploids often differ in physiological and life-history characteristics that may confer adaptive advantages compared with their diploid ancestors. Previous studies have also revealed a causal relationship between polyploidy and population expansion or ecological radiation in plants [63-65]. For example, Ramsey [66] compared the adaptability to the environment for *Achillea borealis* with different ploidy and found that the adaptability to the environment of hexaploid *A. borealis* was stronger than that of tetraploid and more than 70% of the fitness advantage was produced in the polyploidization process. In the present study, the polyploid lineages (group 2 and group 3) showed a historical population expansion (Additional file 1: Figure S3), but the population expansion was not observed in diploid populations (group 2). The results of niche overlap and identity tests between three main ploidy (2x, 4x, and 6x) indicate that the environmental niche of hexaploid var*. deliciosa* was distinguishable from both diploid var*. chinensis* and tetraploid *A. chinensis* complex (Additional file 1: Figure S4). In particular, the hexaploid var*. deliciosa* colonized westward to the areas with cooler and slightly drier environments compared with their diploid parents (Figure 2). Although the ecological niche of tetraploid overlaps with that of diploid, tetraploid usually occupied higher altitudes in the sympatric region. This ecological

radiation pattern is revealed in other plants such as *Aster amellus* [67], *Larrea tridentate* [68], or *Senecio carniolicus* [69] where higher ploidy cytotypes could experience niche expansion while low levels were unable to occupy their full potential niche [70]. Thus, polyploidization played an important role in the historical demography of the *A. chinensis* complex through improving their adaptability to environmental changes.

Conclusions

In the present study, we collected the samples of the *A. chinensis* complex across its distribution ranges in subtropical China and conducted studies on the cytotype distribution and the phylogeography of the species complex. The most frequent and widespread cytotypes were diploids followed by tetraploids and hexaploids, while triploids and octaploids were only observed in a few populations. The ENMs and BBM results suggested that Pliocene and Plio-Pleistocene climatic fluctuations and vicariance may have played key roles in shaping the current population structure and historical demography of the *A. chinensis* complex. This is consistent with the hypothesis that tracking suitable habitat by expansion or contraction would be the main pathway responding to climate changes for the species with a conservative ecological niche. Besides, the polyploidization of the *A. chinensis* complex also played an important role in the historical demography of adaptation to environmental changes.

Methods

Plant sampling

From 2015-2018, 3596 samples were collected from 128 populations representing the species complex *A. chinensis* Planch. including two main species of var. *chinensis* and var. *deliciosa*. Another variety, *A. chinensis* var. *setosa*, narrowly distributed in the Ali mountain of Taiwan island was not included in the present study due to be lack of sample. The field collection followed the ethics and legality of the local government and was permitted by the government. The formal identification of the plant material was undertaken by the National *Actinidia* Germplasm Repository (NAGR) of China, and voucher specimens were also deposited at NAGR. Each population consisted of either var. *chinensis* or var. *deliciosa*, and there were still some populations containing both varieties based on morphology in the sympatric regions. For each population, 6 to 30 individuals were sampled (Additional file 2: Table S1). To avoid collecting closely related individuals, sampled individuals were spaced by ca. 50-100 m. Fresh leaves were dried in silica gel in the field and then stored at -80 °C. In addition, one-year-old shoots were stored at 4 °C in a refrigerator at least 30 days to overcome dormancy and then placed in fresh water at room temperature to stimulate bud breakage for the collection of fresh leaf tissue. The latitude, longitude, and altitude of each sampled individuals were recorded using a global positioning system.

Ploidy analysis

The ploidy level of each individual was determined by estimating its relative DNA content using FCM. Briefly, fresh leaves were chopped in 0.5 mL of nuclear extraction buffer and were then filtered through a nylon sieve, followed by the addition of 2 mL 6-dianidino-2-phenylindole (DAPI) held at 4 °C. Two minutes later, these samples were used to estimate the ploidy levels on CyFlow Ploidy Analyser (Partec, Germany) automatically. Relative fluorescence intensity of a diploid var*. chinensis* cultivar, 'Hongyang' (2n = 58), whose chromosome number had been determined previously based on chromosome counts and the estimated genome size of 750 Mb [71], was used as an internal standard for each measurement [24].

DNA Isolation, Amplification, and Sequencing

For chloroplast DNA analysis, 49 out of 128 populations were selected to cover the entire distribution ranges of the *A. chinensis* complex. A modified CTAB method [72] was used to extract total genomic DNA from 608 individuals in this pool, then 1% agarose gels and NanoDrop 8000 (Thermo Fisher Scientific, Waltham, MA, USA) was employed to examine the quality and concentration of genomic DNA, respectively.

To identify chloroplast DNA (cpDNA) regions with sufficient variation, 12 individuals of var*. chinensis* and 12 var*. deliciosa* were randomly selected to undertake the preliminary screening of primer pairs for eight intron regions: *ndh*F-*rpl*32*, ps*16-*trn*Q*, trn*E-*trn*T*, psb*A-*trn*H*, rpl*16*, trn*L-*trn*F*, trn*L-*trn*T*, and trn*D-*trn*T (data not shown). Finally, three consistently amplified and variable non-coding intergenic spacer (IGS) regions *ndh*F-*rpl*32*, ps*16-*trn*Q, and *trn*E-*trn*T [73](Additional file 2: Table S5) of the plastid genome were selected for further analysis. A total of 608 individuals from all 49 populations of the *A. chinensis* complex were sequenced. Amplification was carried out in a volume of 20 μL reaction solution containing 10 μL 2 × Taq PCR MasterMix (Bioteke), 0.5 μL each primer (0.2 μM), 1 μL template DNA (ca. 50-100 ng) and 8 μL ddH2O. The thermo-cycling conditions for PCR are given as follows: 94 °C, 4 min; 35 × (94 °C, 30 s; 56 °C, 60 s; 72 °C, 60 s); and 72 °C, 10 min. PCR products were sequenced in an ABI 377XL DNA sequencer (Applied Biosystems). All haplotype sequences identified in this study were submitted to GenBank (MT812986-MT813020). Besides, for subsequent analysis, *A. eriantha* and *A. polygama* were used as outgroup and these chloroplast sequences were downloaded from GenBank (KY100979.1 and KX345297.1).

Analyses of genetic diversity and population structure

Three cpDNA fragments were aligned separately and trimmed in MEGA 6 [74] and combined into a single dataset using an online tool FaBox [75] for the phylogeographic analyses. During the analyses, indels (gaps) were treated as a single mutation event and coded as substitutions A or T (a third G or C will be used when three kinds of mutations coexisted). The number of cpDNA haplotypes, haplotype diversities (H_d), and nucleotide diversities (π) [76] for each variety were calculated using DnaSP 5.10 [77]. Haplotype distribution maps were constructed using ArcGIS 10.3. Total genetic diversity (H_T) and within-population diversity (H_S) were calculated with PermutCpSSR 2.0 [78].

The phylogeographic structure was inferred by comparing population differentiation for phylogenetically ordered (N_{ST}) and unordered (G_{ST}) haplotype using PermutCpSSR 2.0. A test of significance comprising 1000 permutations was used to determine if $N_{ST} > G_{ST}$. A significantly higher N_{ST} than G_{ST} usually indicates the presence of phylogeographical structure. The geographically and genetically distinguishable groups and the potential barriers between groups were analyzed using SAMOVA 2.0 [79]. Various SAMOVA were run, increasing the number of K groups until the percentage of explained variance among groups reached a limit (K = 2–20). In addition, a molecular variance (AMOVA) analysis was performed using Arlequin v3.5 [80] with significance tested using 1023 permutations to test genetic differentiation among populations and between varieties. Finally, estimates of pairwise $F_{ST}/(1 - F_{ST})$ were regressed against the pairwise natural logarithm of the geographic distance using a Mantel test with 999 random permutations in GENALEX 6.5 to test IBD patterns among the populations [81, 82]. Measures of pairwise F_{ST} distances were calculated using Arlequin, and the pairwise natural logarithms of geographic distance were calculated by GENALEX 6.5.

Haplotype relationships and divergence time estimation

To determine phylogenetic relationships among haplotypes, median-joining networks were constructed using NETWORK v5.0 [83]. The haplotypes of *A. eriantha* and *A. polygama* were also included in constructing haplotype network as well as in molecular dating analysis. we a Bayesian phylogenetic inference on all haplotypes using Markov chain Monte Carlo (MCMC) methods and a strict clock model in BEAST2 v2.5.0 were performed to estimate divergence time among lineages [84]. Meanwhile, a GRT model was selected based on the result of AIC from jModelTest v 2.1.10 test [85]. As a tree prior, the Yule model was specified. Two calibration points from previous studies of *Actinidia* [86] were used to constrain the nodes with a normal distribution prior. One point is the estimated split between *A. polygama* and the other two species (19 Myr, 95%CI: 14-24), and another is the estimated split between *A. eriantha* and *A. chinensis* (11 Myr, 95%CI: 4.7-17.3). The Monte Carlo Markov chain runs were performed every 1×10^9 generations, with sampling every 1×10^5 generations, following a burn-in of the initial 10% cycles. the convergence and effective sample size (ESS) > 200 for all parameters were assessed using Tracer version 1.7.1. After discarding the first 25% trees as burn-in, the rest of the trees were summarized in a maximum clade credibility tree with treeAnnotator version 2.6.0. Finally, the maximum clade credibility tree was visualized in FigTree version 1.4.3 (available at http://tree.bio.ed.ac.uk/ software/ figtree/).

Demographic history analyses and ancestral geographical area reconstruction

All populations were divided into four groups which represent the eastern 2x populations, eastern 4x populations, central-western 2x populations, and central-western 4x-6x populations, respectively. The mismatch distribution analysis using Arlequin 3.5 [80] was conducted to test whether each group has undergone recent demographic or spatial population expansion events. Multimodal mismatch distributions of pairwise differences between individuals are expected for populations at demographic equilibrium with a relatively stable size over time, whereas unimodal distributions are expected for the population that has experienced recent demographic expansions [87]. The goodness-of-fit of observed

mismatch distributions to the theoretical distributions under a model of sudden expansion was tested with the raggedness index [88]. The significance of the raggedness index was obtained by examining the null distribution of 1000 coalescent simulations of these statistics. Small values of the raggedness index suggest sudden expansion, whereas high values of the raggedness index suggest stationary or genetic bottleneck. Second, for expanding lineage, the expansion parameter (τ) and its 95% confidence interval were converted into generation time (T) since expansion following the equation: $T = \tau/2\mu$ [89], where μ is the neutral mutation rate of the entire cpDNA sequences per generation. The value for μ was calculated as $\mu = uk$, where u is substitution rate in substitutions/site/generation (s/s/g) (here, 1.0×10^{-8}; refer to Gaut [90]), k is the average sequence length of the cpDNA region under this study (here, 1,854 bp), the expansion time was calculated assuming a generation time of 7 years for *A. chinensis* under natural conditions [86]. Finally, we also calculated Tajima's *D* [91] and F_S statistics of F_u [92] using Arlequin 3.5 [80], which is based on 1000 random permutations. The mismatch distribution of the observed number of nucleotide differences between pairs of DNA sequences was computed using DnaSP 5.10 [77].

Ancestral range reconstruction was conducted to estimate possible historical patterns of geographical distribution for *A. chinensis* complex using the BBM (Bayesian binary MCMC) method implemented in RASP v.4.0 [93]. For BBM analysis, 10001 BEAST-generated trees and consensus tree without outgroups were used as topology input, excluding the first 1000 as burning. In the ancestral area reconstruction, the geographical areas of the *A. chinensis* complex were defined four biogeographical regions: EA, East; SO, South; CE, Central; and WE, West (Figure 4a) [54]. Ten BBM chains were run for 100,000 generations with a sampling frequency of 100. F81 was used as a state frequencies model according to the Akaike Information Criterion.

Ecological niche models

To study the niche differentiation and distributional changes for each variety, the Maxent Version 3.4.0 program [94, 95] was used to undertake to test of the ecological niche models and projected their potential distributions during three periods: the present, the Last Glacial Maximum (LGM, 21kya), and the Last Interglacial (LIG,120-140 kya). The distribution data were obtained from our field collection and the Chinese Virtual Herbarium (CVH, available at http://www.cvh.ac.cn/). For all CVH records, the taxonomic identity was verified by specimen images and removed the artificial cultivation records to ensure that geolocations were consistent with the known distribution ranges. After removing duplicates, 150 sites for var*. chinensis* and 90 sites for var*. deliciosa* were identified. The current and past environmental datasets of 19 bioclimate variables with spatial resolutions of 2.5 arc minute were downloaded from the WorldClim database [96]. To avoid over-fitting [97], strongly correlated bioclimatic parameters, according to Pearson's coefficient ($|r| > 0.8$) using the Perl program ENMTools v.1.3 [98], were excluded. Finally, six bioclimatic variables (Additional file 2: Table S6) were retained for ecological niche model construction with Maxent Version 3.4.0. And the model quality was assessed by cross-validation comprising 100 replicates using 25% of the data for model testing, and the accuracy of each cross-validation test was evaluated using the area under the ROC curve (AUC) [99].

Ecology divergence analysis

To measure niche divergence, the environmental niches of the two varieties were compared using niche overlap and identity tests, in which the difference between the actual niches was contrasted with null models generated from randomly reshuffled occurrence points [100]. Niche identity was calculated using Schoener's *D* similarity index [101] and Warren's *I* [100] implemented in ENMTools version 1.3 [98, 100]. Both Schoener's *D* and Warren's *I* ranged from 0 (no niche overlap) to 1 (identical niches). One hundred pseudoreplicates of shuffling were conducted to generate null models for these statistics, and tested for significance; histograms were drawn using R 3.6.0 (http://www.rproject.org/). In addition, the same tests and analyses were also carried out for the three main different level ploidy (2x, 4x, and 6x) populations.

To better understand and visualize the similarities and differences between the environments in which var. *chinensis* and var. *deliciosa* occur, the principal components analyses (PCA) were conducted using the elevation and 19 bioclimatic variables from WorldClim [96, 102]. The same 240 occurrence data points used for ENMs have been extracted values in ArcGIS 10.3. And the matrices of environmental values were standardized before the PCAs were performed, simultaneously the distributions of species in environmental space were visualized in R 3.6.0.

Abbreviations

ENMs
ecological niche models
LIG
Last Interglacial
LGM
Last Glacial Maximum
FCM
flow cytometric measurement
MDA
mismatch distribution analysis
SSD
Sum of square deviations
HRI
Harpending's raggedness index
IBD
isolation-by-distance
RASP
Reconstruct Ancestral State in Phylogenies
BBM
Bayesian Binary MCMC
AUC

Area Under the Curve

LDD

long-distance dispersal

NAGR

the National *Actinidia* Germplasm Repository

DAPI

6-dianidino-2-phenylindole

CVH

Chinese Virtual Herbarium

MCMC

Markov chain Monte Carlo

PCA

principal components analyses

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All sequences of *A. chinensis* complex generated during the current study are available in the NCBI GenBank database under accession numbers MT812986-MT813020.

Competing interests

The authors declare that there are no competing interests.

Funding

The work was supported by grants from the National Key Research and Development Program of China (2019YFD1000200) and the National Natural Science Foundation of China (31370251 and 31770374). These funding bodies did not have a role in the design of the study, the collection, analysis, and interpretation of data, or in writing the manuscript.

Authors' contributions

XHY and ZZL conceived the study and designed the experiments; ZZL, CLY and ZW collected the field samples; CLY and ZW carried out the experiments and analyzed the data; ZW, CHZ, DWL, ZZL and XHY

drafted the manuscript and carried out the critical revision of the manuscript. All authors have read and approved the final manuscript.

Acknowledgments

Not applicable.

Author information

Affiliations

CAS Key Laboratory of Plant Germplasm Enhancement and Speciality Agriculture, Wuhan Botanical Garden, the Chinese Academy of Sciences, Wuhan 430074, Hubei, China.

Zhi Wang, Caihong Zhong, Dawei Li, Chunlin Yan, Zuozhou Li, Xiaohong Yao

College of Life Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China.

Zhi Wang

References

1. Qiu YX, Fu CX, Comes HP. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol Phylogenet Evol. 2011;59(1):225–44.
2. Liu J, Moller M, Provan J, Gao LM, Poudel RC, Li DZ. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. New Phytol. 2013;199(4):1093–108.
3. Khan G, Zhang F, Gao Q, Fu P, Zhang Y, Chen S. Spiroides shrubs on Qinghai-Tibetan Plateau: Multilocus phylogeography and palaeodistributional reconstruction of *Spiraea alpina* and S. Mongolica (Rosaceae). Mol Phylogenet Evol. 2018;123:137 – 48.
4. Fu PC, Sun SS, Khan G, Dong XX, Tan JZ, Favre A, Zhang FQ, Chen SL. Population subdivision and hybridization in a species complex of *Gentiana* in the Qinghai-Tibetan Plateau. Ann Bot. 2020;125(4):677–90.
5. Yu XQ, Gao LM, Soltis DE, Soltis PS, Yang JB, Fang L, Yang SX, Li DZ. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. New Phytol. 2017;215(3):1235–48.
6. Fan D, Sun Z, Li B, Kou Y, Hodel RGJ, Jin Z, Zhang Z. Dispersal corridors for plant species in the Poyang Lake Basin of southeast China identified by integration of phylogeographic and geospatial data. Ecol Evol. 2017;7(14):5140–8.
7. Chen YS, Deng T, Zhou Z, Sun H. Is the East Asian flora ancient or not? Nat Sci Rev. 2018;5(6):920–32.

8. Zhang J, Li Z, Fritsch PW, Tian H, Yang A, Yao X. Phylogeography and genetic structure of a Tertiary relict tree species, *Tapiscia sinensis* (Tapisciaceae): implications for conservation. Ann Bot. 2015;116(5):727–37.
9. Zhang YH, Wang IJ, Comes HP, Peng H, Qiu YX. Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure in a Tertiary relict species, *Emmenopterys henryi* (Rubiaceae). Sci Rep. 2016;6:24041.
10. Riddle BR. Comparative phylogeography clarifies the complexity and problems of continental distribution that drove A. R. Wallace to favor islands. Proc Natl Acad Sci U S A. 2016;113(29):7970–7.
11. Kou Y, Zhang L, Fan D, Cheng S, Li D, Hodel RGJ, Zhang Z. Evolutionary history of a relict conifer, *Pseudotaxus chienii* (Taxaceae), in south-east China during the late Neogene: old lineage, young populations. Ann Bot. 2020;125(1):105–17.
12. Gong W, Liu W, Gu L, Kaneko S, Koch MA, Zhang D. From glacial refugia to wide distribution range: demographic expansion of *Loropetalum chinense* (Hamamelidaceae) in Chinese subtropical evergreen broadleaved forest. Org Divers Evol. 2016;16(1):23–38.
13. Gao YD, Gao XF, Harris A. Species Boundaries and Parapatric Speciation in the Complex of Alpine Shrubs, *Rosa sericea* (Rosaceae), Based on Population Genetics and Ecological Tolerances. Front Plant Sci. 2019;10:321.
14. Soltis PS, Marchant DB, Van de Peer Y, Soltis DE. Polyploidy and genome evolution in plants. Curr Opin Genet Dev. 2015;35:119–25.
15. Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. The frequency of polyploid speciation in vascular plants. Proc Natl Acad Sci U S A. 2009;106(33):13875–9.
16. Husband BC, Baldwin SJ, Suda J. The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. In: Plant genome diversity volume 2. Springer; 2013. p. 255 – 76.
17. Wu LL, Cui XK, Milne RI, Sun YS, Liu JQ. Multiple autopolyploidizations and range expansion of *Allium przewalskianum* Regel. (Alliaceae) in the Qinghai-Tibetan Plateau. Mol Ecol. 2010;19(8):1691–704.
18. He J, Wang S, Li J, Fan Z, Liu X, Wang Y. Genetic differentiation and spatiotemporal history of diploidy and tetraploidy of *Clintonia udensis*. Ecol Evol. 2017;7(23):10243–51.
19. Zhang X, Su H, Yang J, Feng L, Li Z, Zhao G. Population genetic structure, migration, and polyploidy origin of a medicinal species *Gynostemma pentaphyllum* (Cucurbitaceae). Ecol Evol. 2019;9(19):11145–70.
20. Li J-Q, Soejarto D, Li X-W. A Revision of the Genus Actinidia from China. In: VI International Symposium on Kiwifruit 753: 2007. 41 – 4.
21. Liu Y, Li D, Yan L, Huang H. The microgeographical patterns of morphological and molecular variation of a mixed ploidy population in the species complex *Actinidia chinensis*. PLoS ONE. 2015;10(2):e0117596.

22. Tian Z. Genetic diversity, gene introgression and homoplasy in sympatric populations of the genus Actinidia as revealed by chloroplast microsatellite markers. Biodiversity Science. 2007;15(1):1.
23. Liu Y, Liu Y, Huang H. Genetic variation and natural hybridization among sympatric *Actinidia* species and the implications for introgression breeding of kiwifruit. Tree Genet Genomes. 2010;6(5):801–13.
24. Li D, Liu Y, Zhong C, Huang H. Morphological and cytotype variation of wild kiwifruit (*Actinidia chinensis* complex) along an altitudinal and longitudinal gradient in central-west China. Bot J Linnean Soc. 2010;164(1):72–83.
25. Ferguson AR, Huang H. Genetic resources of kiwifruit: domestication and breeding. Horticultural Reviews. 2007;33:1–121.
26. Wang Y-C, Liao L, Li Z-Z. Genetic differentiation of *Actinidia chinensis* and analysis of gene flow barriers in the Qinling Mountains, the species' northern distribution boundary. Genet Resour Crop Evol. 2017;65(3):881–95.
27. Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodriguez-Sanchez F, Dobrowski SZ, Hampe A, Hu FS, Ashcroft MB, Bartlein PJ, et al. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. New Phytol. 2014;204(1):37–54.
28. Wang Y-H, Jiang W-M, Comes HP, Hu FS, Qiu Y-X, Fu C-X. Molecular phylogeography and ecological niche modelling of a widespread herbaceous climber, *Tetrastigma hemsleyanum* (Vitaceae): insights into Plio-Pleistocene range dynamics of evergreen forest in subtropical China. New Phytol. 2015;206(2):852–67.
29. Scoble J, Lowe AJ. A case for incorporating phylogeography and landscape genetics into species distribution modelling approaches to improve climate adaptation and conservation planning. Divers Distrib. 2010;16(3):343–53.
30. Mellick R, Lowe A, Allen C, Hill RS, Rossetto M. Palaeodistribution modelling and genetic evidence highlight differential post-glacial range shifts of a rain forest conifer distributed across a latitudinal gradient. J Biogeogr. 2012;39(12):2292–302.
31. Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science. 1983;220(4601):1049–51.
32. Weiss H, DobeS C, Schneeweiss GM, Greimler J. Occurrence of tetraploid and hexaploid cytotypes between and within populations in *Dianthus* sect. *Plumaria* (Caryophyllaceae). New Phytol. 2002;85–94.
33. Levin DA. Polyploidy and novelty in flowering plants. Am Nat. 1983;122(1):1–25.
34. Atkinson RG, Cipriani G, Whittaker DJ, Gardner RC. The allopolyploid origin of kiwifruit, *Actinidia deliciosa* (Actinidiaceae). Plant Syst Evol. 1997;205(1–2):111–24.
35. Huang H, Dane F, Wang Z, Jiang Z, Huang R, Wang S. Isozyme inheritance and variation in *Actindia*. Heredity. 1997;78(3):328–36.
36. Kolar F, Stech M, Travnicek P, Rauchova J, Urfus T, Vit P, Kubesova M, Suda J. Towards resolving the *Knautia arvensis* agg. (Dipsacaceae) puzzle: primary and secondary contact zones and ploidy segregation at landscape and microgeographic scales. Ann Bot. 2009;103(6):963–74.

37. Sabara HA, Kron P, Husband BC. Cytotype coexistence leads to triploid hybrid production in a diploid-tetraploid contact zone of *Chamerion angustifolium* (Onagraceae). Am J Bot. 2013;100(5):962–70.
38. Mraz P, Spaniel S, Keller A, Bowmann G, Farkas A, Singliarova B, Rohr RP, Broennimann O, Muller-Scharer H. Anthropogenic disturbance as a driver of microspatial and microhabitat segregation of cytotypes of *Centaurea stoebe* and cytotype interactions in secondary contact zones. Ann Bot. 2012;110(3):615–27.
39. Zozomova-Lihova J, Malanova-Krasna I, Vit P, Urfus T, Senko D, Svitok M, Kempa M, Marhold K. Cytotype distribution patterns, ecological differentiation, and genetic structure in a diploid-tetraploid contact zone of *Cardamine amara*. Am J Bot. 2015;102(8):1380–95.
40. Hulber K, Sonnleitner M, Suda J, Krejcikova J, Schonswetter P, Schneeweiss GM, Winkler M. Ecological differentiation, lack of hybrids involving diploids, and asymmetric gene flow between polyploids in narrow contact zones of *Senecio carniolicus* (syn. *Jacobaea carniolica*, Asteraceae). Ecol Evol. 2015;5(6):1224–34.
41. Winkler M, Escobar Garcia P, Gattringer A, Sonnleitner M, Hulber K, Schonswetter P, Schneeweiss GM. A novel method to infer the origin of polyploids from Amplified Fragment Length Polymorphism data reveals that the alpine polyploid complex of *Senecio carniolicus* (Asteraceae) evolved mainly via autopolyploidy. Mol Ecol Resour. 2017;17(5):877–92.
42. Arnold B, Kim S-T, Bomblies K. Single Geographic Origin of a Widespread Autotetraploid *Arabidopsis arenosa* Lineage Followed by Interploidy Admixture. Mol Biol Evol. 2015;32(6):1382–95.
43. Mandak B, Vit P, Krak K, Travnicek P, Havrdova A, Hadincova V, Zakravsky P, Jarolimova V, Bacles CF, Douda J. Flow cytometry, microsatellites and niche models reveal the origins and geographical structure of *Alnus glutinosa* populations in Europe. Ann Bot. 2016;117(1):107–20.
44. Qi XS, Chen C, Comes HP, Sakaguchi S, Liu YH, Tanaka N, Sakio H, Qiu YX. Molecular data and ecological niche modelling reveal a highly dynamic evolutionary history of the East Asian Tertiary relict *Cercidiphyllum* (Cercidiphyllaceae). New Phytol. 2012;196(2):617–30.
45. Chen J-M, Zhao S-Y, Liao Y-Y, Gichira AW, Gituru RW, Wang Q-F. Chloroplast DNA phylogeographic analysis reveals significant spatial genetic structure of the relictual tree *Davidia involucrata* (Davidiaceae). Conserv Genet. 2015;16(3):583–93.
46. Cao Y-N, Comes HP, Sakaguchi S, Chen L-Y, Qiu Y-X. Evolution of East Asia's Arcto-Tertiary relict *Euptelea* (Eupteleaceae) shaped by Late Neogene vicariance and Quaternary climate change. BMC Evol Biol. 2016;16.
47. Hewitt GM. Speciation, hybrid zones and phylogeography or seeing genes in space and time. Mol Ecol. 2001;10(3):537–49.
48. Heermance RV, Chen J, Burbank DW, Wang C. Chronology and tectonic controls of Late Tertiary deposition in the southwestern Tian Shan foreland, NW China. Basin Res. 2007;19(4):599–632.
49. Ge J, Dai Y, Zhang Z, Zhao D, Li Q, Zhang Y, Yi L, Wu H, Oldfield F, Guo Z. Major changes in East Asian climate in the mid-Pliocene: Triggered by the uplift of the Tibetan Plateau or global cooling? J Asian Earth Sci. 2013;69:48–59.

50. Shi Y, Li J, Li B. Uplift of the Qinghai-Xizang (Tibetan) plateau and east Asia environmental change during late Cenozoic. ACTA GEOGRAPHICA SINICA-CHINESE EDITION. 1999;54:20–8.
51. Yan HF, Zhang CY, Wang FY, Hu CM, Ge XJ, Hao G. Population expanding with the phalanx model and lineages split by environmental heterogeneity: a case study of *Primula obconica* in subtropical China. PLoS One. 2012;7(9):e41315.
52. Kou Y, Cheng S, Tian S, Li B, Fan D, Chen Y, Soltis DE, Soltis PS, Zhang Z. The antiquity of *Cyclocarya paliurus* (Juglandaceae) provides new insights into the evolution of relict plants in subtropical China since the late Early Miocene. J Biogeogr. 2016;43(2):351–60.
53. Zhang X-W, Li Y, Zhang Q, Fang Y-M. Ancient east-west divergence, recent admixture, and multiple marginal refugia shape genetic structure of a widespread oak species (*Quercus acutissima*) in China. Tree Genet Genomes. 2018;14(6).
54. Huang H, Liu Y. Natural hybridization, introgression breeding, and cultivar improvement in the genus *Actinidia*. Tree Genet Genomes. 2014;10(5):1113–22.
55. Tian S, Pujol JLp, Wang HW, Ge S, Zhang Z-Y. Molecular evidence for glacial expansion and interglacial retreat during Quaternary climatic changes in a montane temperate pine (*Pinus kwangtungensis* Chun ex Tsiang) in southern China. Plant Syst Evol. 2010;284(3–4): 219 – 29.
56. Davis MB, Shaw RG. Range shifts and adaptive responses to Quaternary climate change. Science. 2001;292(5517):673–9.
57. Colwell RK, Brehm G, Cardelús CL, Gilman AC, Longino JT. Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. Science. 2008;322(5899):258–61.
58. Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. Science. 2006;311(5768):1751–3.
59. Visger CJ, Germain-Aubrey CC, Patel M, Sessa EB, Soltis PS, Soltis DE. Niche divergence between diploid and autotetraploid *Tolmiea*. Am J Bot. 2016;103(8):1396–406.
60. Munoz-Pajares AJ, Perfectti F, Loureiro J, Abdelaziz M, Biella P, Castro M, Castro S, Gomez JM. Niche differences may explain the geographic distribution of cytotypes in *Erysimum mediohispanicum*. Plant Biol. 2018;20:139–47.
61. Lopez-Jurado J, Mateos-Naranjo E, Balao F. Niche divergence and limits to expansion in the high polyploid *Dianthus broteri* complex. New Phytol. 2019;222(2):1076–87.
62. Molina-Henao YF, Hopkins R. Autopolyploid lineage shows climatic niche expansion but not divergence in *Arabidopsis arenosa*. Am J Bot. 2019;106(1):61–70.
63. Leitch A, Leitch I. Genomic plasticity and the diversity of polyploid plants. Science. 2008;320(5875):481–3.
64. Van de Peer Y, Mizrachi E, Marchal K. The evolutionary significance of polyploidy. Nat Rev Genet. 2017;18(7):411.
65. Han TS, Zheng QJ, Onstein RE, Rojas-Andrés BM, Hauenschild F, Muellner-Riehl AN, Xing YW. Polyploidy promotes species diversification of *Allium* through ecological shifts. New Phytol.

2020;225(1):571–83.
66. Ramsey J. Polyploidy and ecological adaptation in wild yarrow. Proc Natl Acad Sci U S A. 2011;108(17):7096–101.
67. Münzbergová Z, Šurinová M, Castro S. Absence of gene flow between diploids and hexaploids of *Aster amellus* at multiple spatial scales. Heredity. 2013;110(2):123–30.
68. Laport RG, Minckley RL, Ramsey J. Ecological distributions, phenological isolation, and genetic structure in sympatric and parapatric populations of the *Larrea tridentata* polyploid complex. Am J Bot. 2016;103(7):1358–74.
69. Sonnleitner M, Hülber K, Flatscher R, García PE, Winkler M, Suda J, Schönswetter P, Schneeweiss GM. Ecological differentiation of diploid and polyploid cytotypes of *Senecio carniolicus sensu lato* (Asteraceae) is stronger in areas of sympatry. Ann Bot. 2016;117(2):269–76.
70. Kolar F, Certner M, Suda J, Schonswetter P, Husband BC. Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. Trends Plant Sci. 2017;22(12):1041–55.
71. Huang S, Ding J, Deng D, Tang W, Sun H, Liu D, Zhang L, Niu X, Zhang X, Meng M, et al. Draft genome of the kiwifruit *Actinidia chinensis*. Nat Commun. 2013;4:2640.
72. Doyle J, Doyle J. Isolation of DNA from small amounts of plant tissues. BRL focus. 1990;12(1):V15.
73. Tang P, Xu Q, Shen R, Yao X. Phylogenetic relationship in *Actinidia* (Actinidiaceae) based on four noncoding chloroplast DNA sequences. Plant Syst Evol. 2019;305(9):787–96.
74. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.
75. Villesen P. FaBox: an online toolbox for fasta sequences. Mol Ecol Notes. 2007;7(6):965–8.
76. Nei M. Molecular Evolutionary Genetics. New York: Columbia University Press; 1987.
77. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451–2.
78. Pons O, Petit R. Measwring and testing genetic differentiation with ordered versus unordered alleles. Genetics. 1996;144(3):1237–45.
79. Dupanloup I, Schneider S, Excoffier L. A simulated annealing approach to define the genetic structure of populations. Mol Ecol. 2002;11(12):2571–81.
80. Excoffier L, 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(3):564–7.
81. Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics. 1997;145(4):1219–28.
82. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics. 2012;28(19):2537–9.
83. Bandelt H-J, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16(1):37–48.

84. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D. De Maio N. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol. 2019;15(4):e1006650.
85. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9(8):772.
86. Liu Y, Li D, Zhang Q, Song C, Zhong C, Zhang X, Wang Y, Yao X, Wang Z, Zeng S, et al. Rapid radiations of both kiwifruit hybrid lineages and their parents shed light on a two-layer mode of species diversification. New Phytol. 2017;215(2):877–90.
87. Harpending H. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum biol. 1994;591–600.
88. Rogers AR, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol. 1992;9(3):552–69.
89. Rogers AR. Genetic evidence for a Pleistocene population explosion. Evolution. 1995;49(4):608–15.
90. Gaut BS. Molecular clocks and nucleotide substitution rates in higher plants. In: Evolutionary biology. Springer; 1998;93–120.
91. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3):585–95.
92. Fu Y-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147(2):915–25.
93. Yu Y, Harris AJ, Blair C, He X. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. Mol Phylogenet Evol. 2015;87:46–9.
94. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190(3–4):231–59.
95. Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 2008;31(2):161–75.
96. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37(12):4302–15.
97. Peterson A, Nakazawa Y. Environmental data sets matter in ecological niche modelling: an example with Solenopsis invicta and Solenopsis richteri. Glob Ecol Biogeogr. 2008;17(1):135–44.
98. Warren DL, Glor RE, Turelli M. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography. 2010;33(3):607–11.
99. Fawcett T. An introduction to ROC analysis. Pattern Recogn Lett. 2006;27(8):861–74.
100. Warren DL, Glor RE, Turelli M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution. 2008;62(11):2868–83.
101. Schoener TW. The Anolis lizards of Bimini: resource partitioning in a complex fauna. Ecology. 1968;49(4):704–26.

102. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005;25(15):1965–78.

Tables

Table 1 Genetic diversity and differentiation analyses for cpDNA variations in var. *chinensis* and var. *deliciosa*

species	H_S	H_T	G_{ST}	N_{ST}
A. chinensis	0.470	0.920	0.490	0.591**
A. deliciosa	0.380	0.552	0.312	0.252
Pooled	0.441	0.884	0.501	0.638**

* Indicates that N_{ST} is significantly different from G_{ST} ($P < 0.01$)

Table 2 Population demography of different groups for *A. chinensis* complex

Group	Parameter(τ)	(*t*,Ma)	*SSD*	*HRI*	Tajima's *D*	Fu's F_S
group 1	5.830 (2.477-10.316)	NC	0.060	0.194	0.920	2.731
group 2	0.446 (0.177-0.724)	0.084 (0.033-0.136)	**0.002***	0.178	**-1.663***	**-3.104***
group 3	0.729 (0.432-0.861)	0.137 (0.081-0.162)	**0.031***	0.215	**-1.230***	-1.408
group 4	3.928 (2.027-6.203)	NC	0.046	0.134	0.484	0.926

Note: *SSD* = Sum of Squared deviation; *HRI* = Harpending's Raggedness index; Tajima's *D* and Fu's F_S are neutral test index; NC, not calculated; *****, $P < 0.05$

Figures

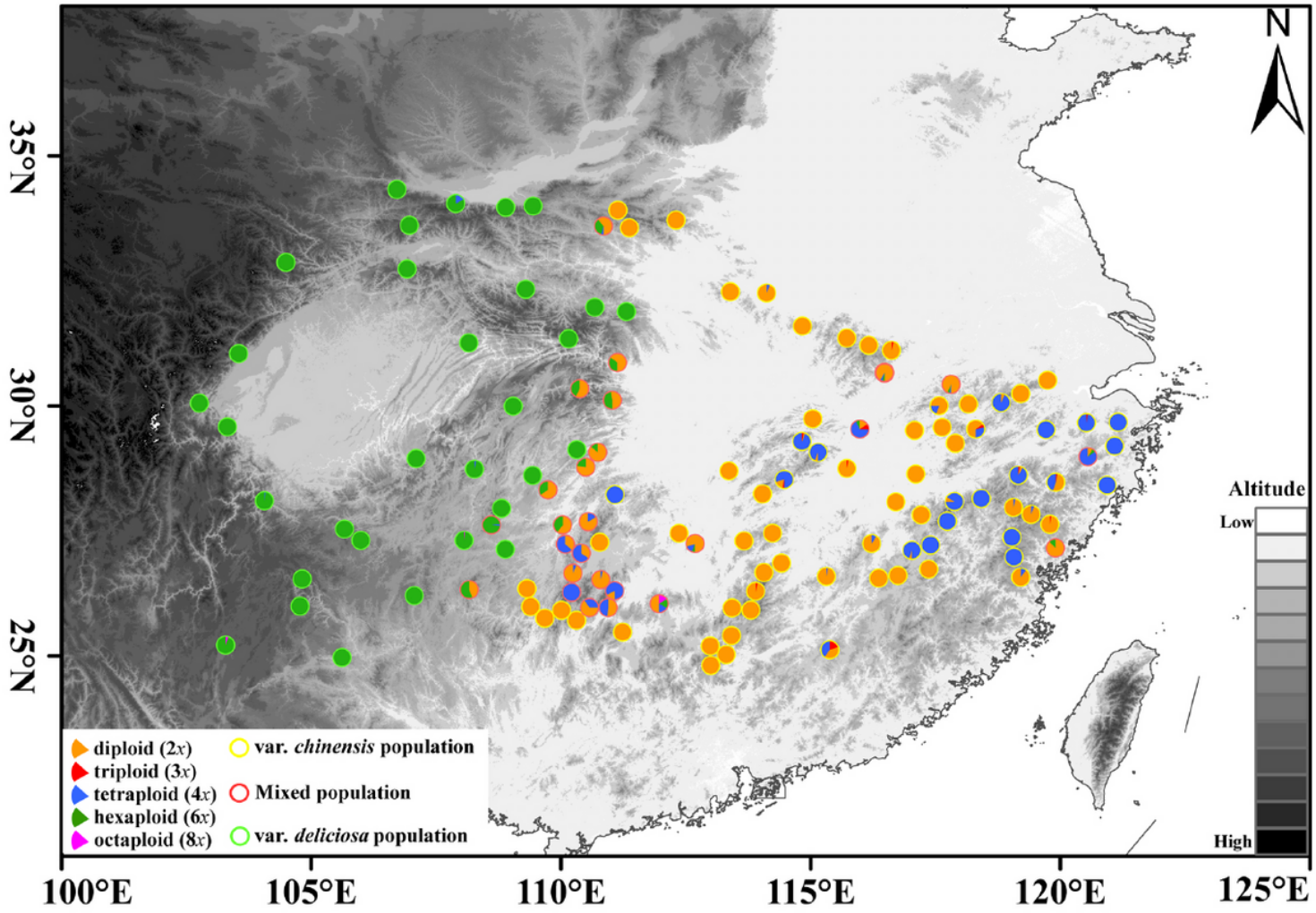

Figure 1

Cytotype distribution of the A. chinensis complex. Each location is represented by a filled circle or pie chart, with ploidy levels indicated by different colors. Map generated in ESRI ArcGIS 10.3. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

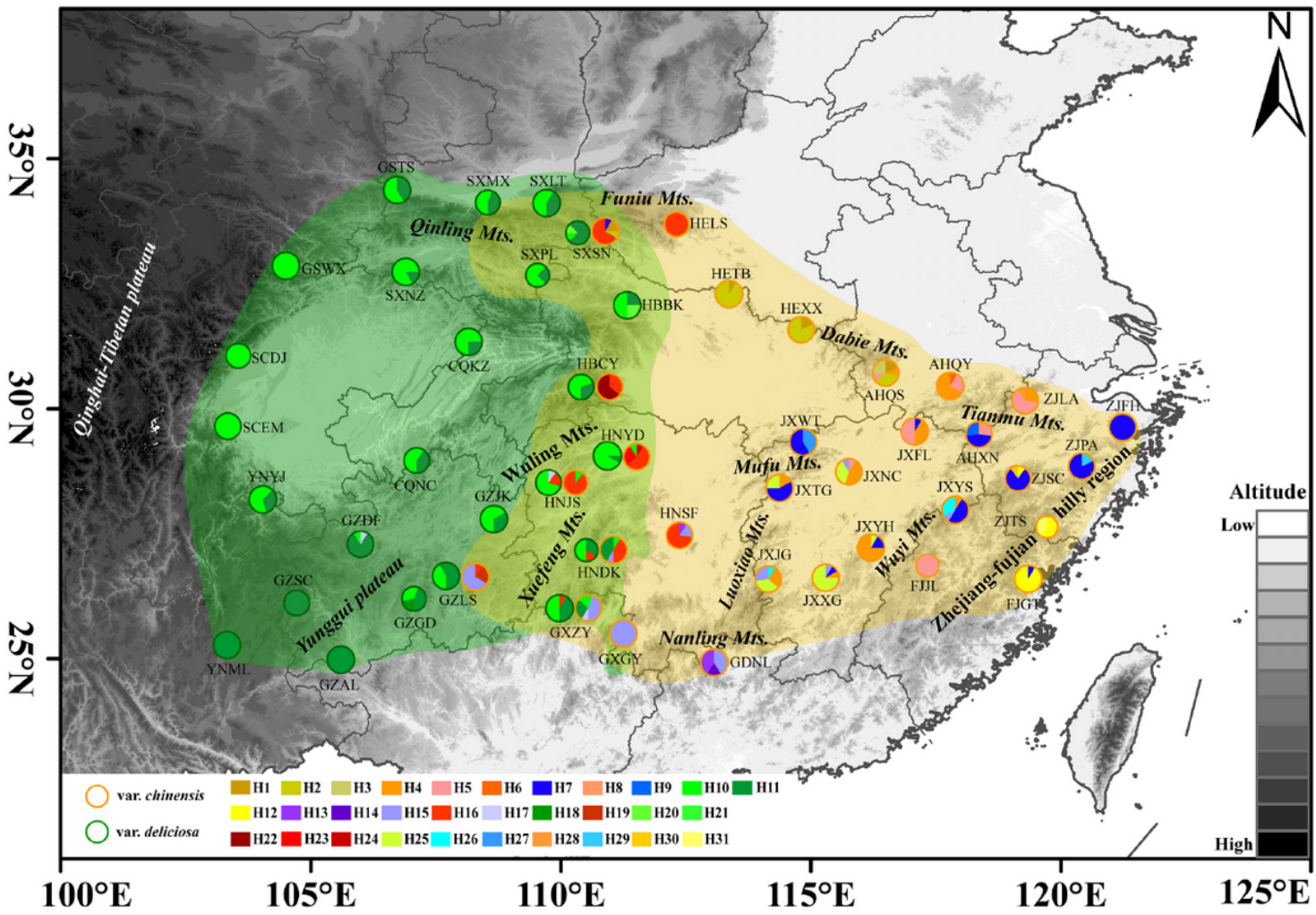

Figure 2

Geographical distribution of cpDNA haplotypes. For the A. chinensis complex, geographical locations of the 49 populations and distributions of 31 chloroplast haplotypes of the A. chinensis complex examined in this study. var. chinensis and var. deliciosa populations were indicated by different colors. Map generated in ESRI ArcGIS 10.3. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

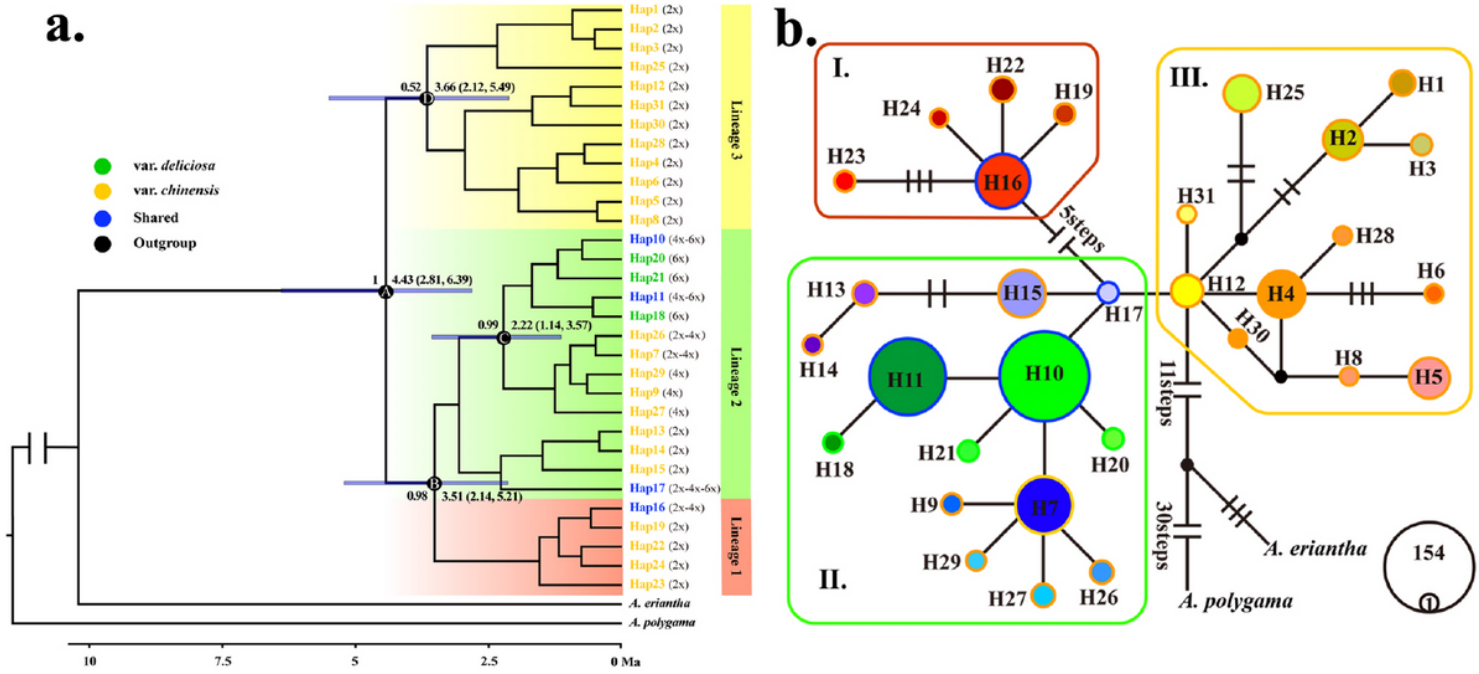

Figure 3

BEAST-derived chronograms and haplotypes network of the A. chinensis complex. (a) BEAST-derived chronograms of the A. chinensis complex based on cp DNA sequences. Blue bars indicate 95% HPD credibility intervals for nodes of particular interest with ages (in Myr ago, Ma). The cytotypes were also labeled for each haplotype. (b)The network of 31 chloroplast haplotypes. Each circle means a unique haplotype, with circle size reflecting its frequency. Small black circles mean missing haplotypes.

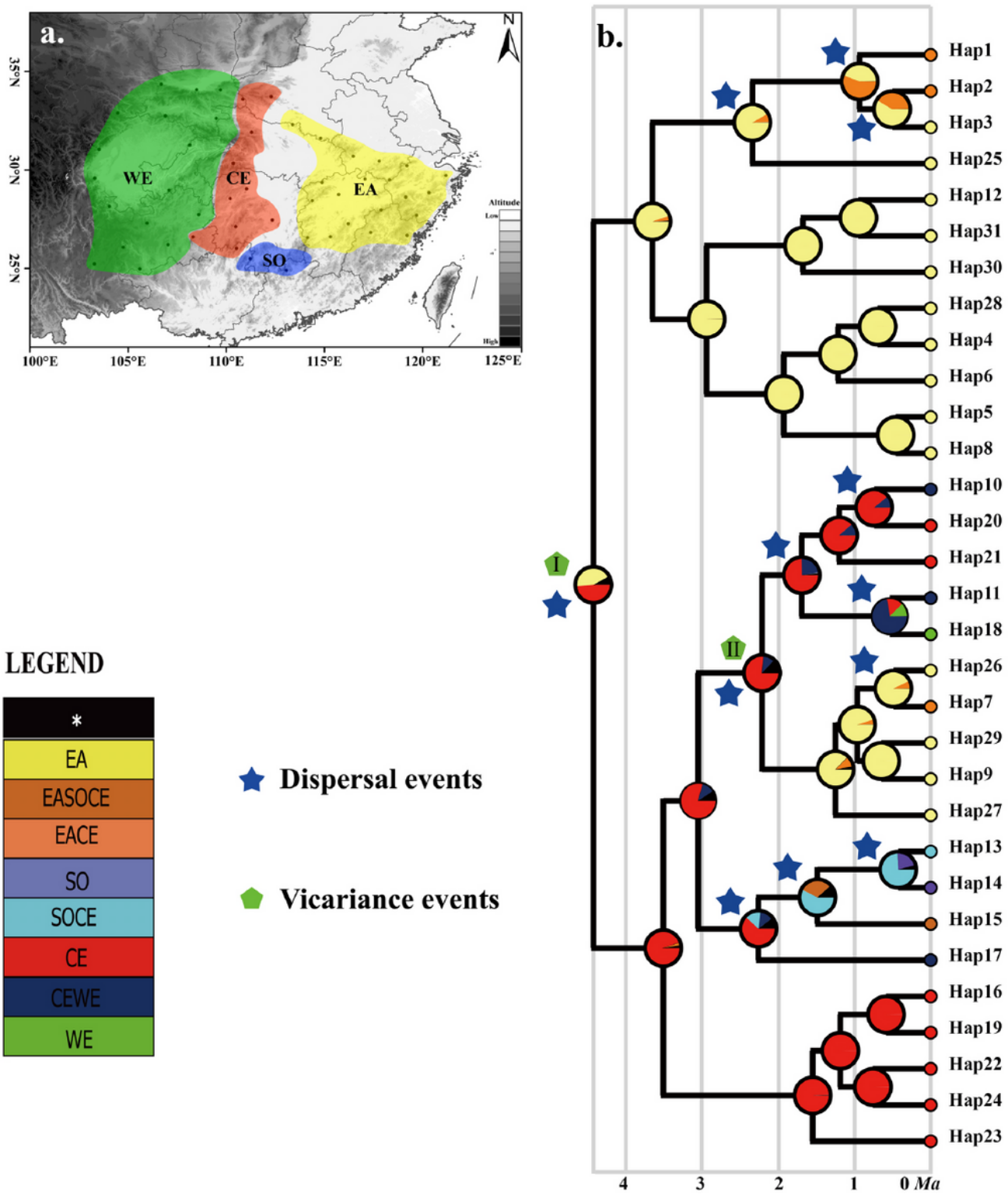

Figure 4

The ancestral areas were reconstructed using the BBM (Bayesian Binary MCMC) method. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

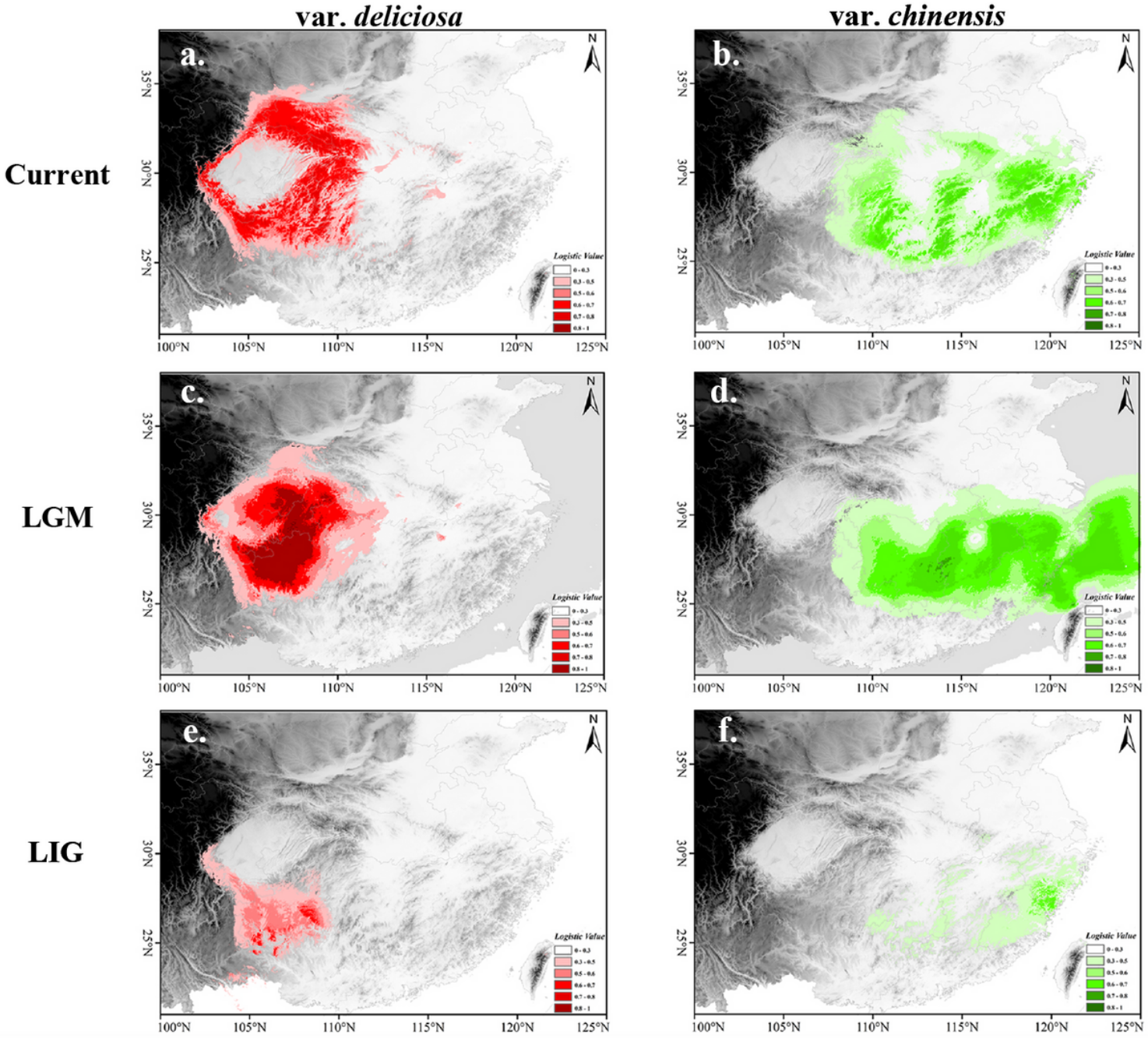

Figure 5

Potential distributions of var. chinensis and var. deliciosa predicted using MaxEnt. Shown at 2.5 arc minute resolution and projected based on six bioclimatic variables representing the current (a,b), LGM (c,d), and LIG (e,f) climatic conditions, respectively. Warmer colors denote areas with a higher probability of presence. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

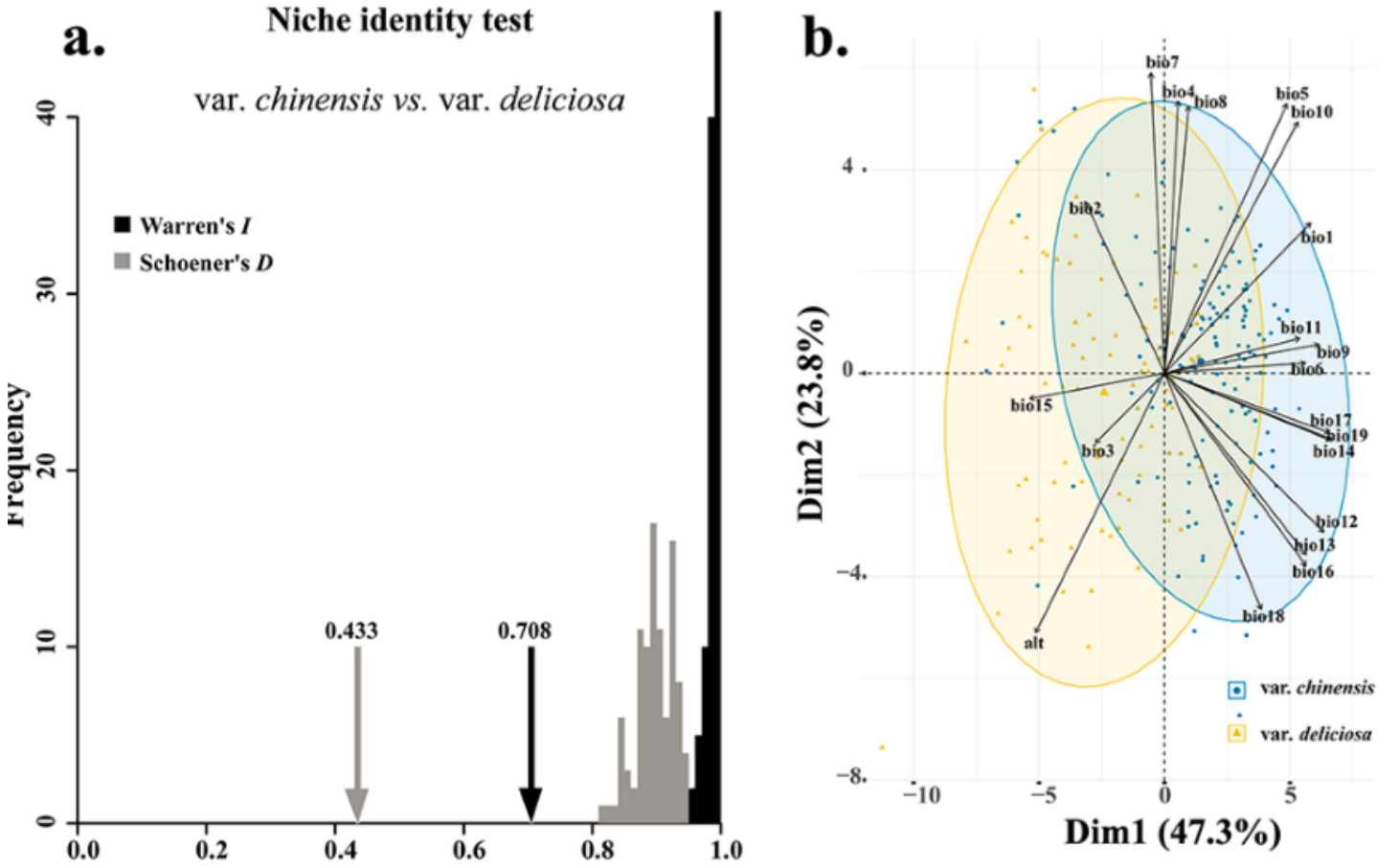

Figure 6

Niche comparation of var. chinensis and var. deliciosa. (a) niche-identity tests by comparing the niches of var. chinensis and var. deliciosa. (b) principal component analysis (PCA) using environmental variables from 240 occurrence data points.

Supplementary Files

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

- Addtionalfile1.pdf
- Addtionalfile2.xls