

RESEARCH

Open Access



Molecular phylogeography and species distribution modelling evidence of 'oceanic' adaptation for *Actinidia eriantha* with a refugium along the oceanic–continental gradient in a biodiversity hotspot

Rui Guo^{1,2,3,4†}, Yong-Hua Zhang^{5†}, Hua-Jie Zhang^{1,2}, Jacob B. Landis^{6,7}, Xu Zhang^{1,2,3}, Heng-Chang Wang^{1,2} and Xiao-Hong Yao^{1,2*}

Abstract

Background: Refugia is considered to be critical for maintaining biodiversity; while discerning the type and pattern of refugia is pivotal for our understanding of evolutionary processes in the context of conservation. Interglacial and glacial refugia have been studied throughout subtropical China. However, studies on refugia along the oceanic–continental gradient have largely been ignored. We used a liana *Actinidia eriantha*, which occurs across the eastern moist evergreen broad-leaved forests of subtropical China, as a case study to test hypotheses of refugia along the oceanic–continental gradient and 'oceanic' adaptation.

Results: The phylogeographic pattern of *A. eriantha* was explored using a combination of three cpDNA markers and 38 nuclear microsatellite loci, Species distribution modelling and dispersal corridors analysis. Our data showed intermediate levels of genetic diversity [haplotype diversity (h_T) = 0.498; unbiased expected heterozygosity (UH_E) = 0.510] both at the species and population level. Microsatellite loci revealed five clusters largely corresponding to geographic regions. Coalescent time of cpDNA lineages was dated to the middle Pliocene (ca. 4.03 Ma). Both geographic distance and climate difference have important roles for intraspecific divergence of the species. The Zhejiang-Fujian Hilly Region was demonstrated to be a refugium along the oceanic–continental gradient of the species and fit the 'refugia in refugia' pattern. Species distribution modelling analysis indicated that Precipitation of Coldest Quarter (importance of 44%), Temperature Seasonality (29%) and Mean Temperature of Wettest Quarter (25%) contributed the most to model development. By checking the isolines in the three climate layers, we found that *A. eriantha* prefer higher precipitation during the coldest quarter, lower seasonal temperature difference and lower mean temperature during the wettest quarter, which correspond to 'oceanic' adaptation. *Actinidia eriantha* expanded to its western distribution range along the dispersal corridor repeatedly during the glacial periods.

*Correspondence: yaoh@wbgcas.cn

†Rui Guo and Yong-Hua Zhang contributed equally to this work.

¹ CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, The Chinese Academy of Sciences, Wuhan 430074, Hubei, China

Full list of author information is available at the end of the article



Conclusions: Overall, our results provide integrated evidence demonstrating that the Zhejiang-Fujian Hilly Region is a refugium along the oceanic–continental gradient of *Actinidia eriantha* in subtropical China and that speciation is attributed to ‘oceanic’ adaptation. This study gives a deeper understanding of the refugia in subtropical China and will contribute to the conservation and utilization of kiwifruit wild resources in the context of climate change.

Keywords: *Actinidia eriantha*, Refugium, Oceanic–continental gradient, ‘Oceanic’ adaptation, Climatic fluctuations, Phylogeography, Subtropical China

Background

Distinguishing the type of refugia is crucial for our understanding of evolutionary processes such as adaptation and speciation, while also forecasting how current climate change may affect the species in the context of conservation [1, 2]. Refugia are habitats for which components of biodiversity retreat to, persist in and can potentially expand from under changing environmental conditions [2]. Refugia are mainly classified into interglacial refugia, glacial refugia and refugia along the oceanic–continental gradient [1]. Interglacial and glacial refugia are usually defined within a latitudinal gradient, while refugia along the oceanic–continental gradient often have a longitudinal perspective [1]. The oceanic–continental gradient in climate provides increasing seasonal range of temperature and decreasing precipitation from the coast to inland areas [3]. Mountain ranges, especially near the coast, intercept clouds and cause major local increase in precipitation [3]. Two types of adaptation have been defined for species with refugia along the oceanic–continental gradient: ‘oceanic’ and ‘continental’. ‘Oceanic’ adaptation implies more humid, less seasonably variable climate; ‘continental’ adaptation consists of a drier climate with greater seasonal variation [1]. Molecular phylogeographic studies at the intraspecific level can reveal patterns of historical demography of species, including areas of refugia [4].

Extensive literature has recently emerged on the phylogeographic patterns of plant species in subtropical China (e.g. [5–10]). Subtropical China is a global biodiversity hotspot, well known due to its species richness, complex topography and fluctuating paleoclimate [11–15]. This area is located between the Qinling Mountains-Huai River line (ca. 34°N) and the Tropic of Cancer (ca. 23°N), bordered by the Qinghai-Tibetan Plateau (ca. 99°E) in the west and the coastline in the east [16]. This region has never been directly covered by ice sheets during glaciation periods of the Pleistocene, thus has preserved numerous Tertiary plant genera [16, 17]. However, the area has undergone complex climate changes during the Quaternary Period. For example, the climate of this region during the LGM (last glacial maxima) was cooler by c. 4–6°C and dryer by c. 400–600 mm/yr [18, 19]. Interglacial and glacial refugia have been

studied thoroughly in subtropical China, shedding light on our understanding of the effects of glacial and post-glacial cycles. For example, glacial refugia were detected for *Lindera aggregata* (Sims) Kosterm [9], *Castanopsis eyrei* (Champ.) Tutch [20], *Machilus thunbergii* Sieb. et Zucc [21], *Loropetalum chinensis* (R. Br.) Oliver [22] and *Sargentodoxa cuneata* (Oliv.) Rehd. et Wils [23] while interglacial refugia were identified for *Emmenopterys henry* Oliv [12]. and *Rosa sericea* Lindl. complex [24]. Large mountains in subtropical China served as important refugia such as the Nanling Mountains for *Lindera aggregata* [9]. The variable topography in these areas offer a large scope of elevational shifts for plant species in response to climatic changes [2]. Despite the gradient of oceanic–continental climate being significantly variable during glacial cycles with important biotic consequences [25], refugia along the oceanic–continental gradient were largely ignored in the area.

Actinidia eriantha Benth. (Ericales: Actinidiaceae; 2n=58) is a suitable model for testing hypotheses of refugia along the oceanic–continental gradient and ‘oceanic’ adaptation, on account of its distribution and moist habitat. The species is an important component of liana species in eastern moist evergreen broad-leaved forests (EBLF) of subtropical China at an altitude of 200 to 1500 m [26]. The more narrow distribution of the species along a series of mountains from the coastline of the East China Sea to the eastern edge of the Yungui Plateau implies an oceanic adaptation. *Actinidia eriantha* is a functionally dioecious, perennial liana species characterized by young branchlets, petioles, inflorescences, sepals and fruits densely tomentose with milky-white to dirty yellow hairs or appressed tomentum [27]. The species is recognized as a valuable species for commercial kiwifruit improvement by extending shelf life and increasing vitamin C concentration, as well as having been used in traditional Chinese medicine to treat gastric carcinoma, nasopharyngeal carcinoma, breast carcinoma, and hepatitis [28]. To assist in collecting germplasm resources and protection of the species, revealing how the species responded to past climate changes and where refugia are located is necessary. To date, a few studies of population genetic differentiation of *A. eriantha* with limited sampling from narrow geographical areas have been

conducted [29–31]. The lack of collections in many areas of the region has prevented the phylogeographic studies for conservation purposes.

Here, we used extensive sampling covering the whole distribution range of the species combined with nuclear SSR and cpDNA evidence, Species Distribution Models (SDM) and dispersal corridors analyses to explore the patterns of historical demography of *A. eriantha* comprehensively. We asked: i) What are the patterns of genetic diversity and population divergence of *A. eriantha*? ii) Which factors have an effect on the genetic structure of the species? iii) What is the location and type of the refugia of the species? iv) What is the direction of dispersal of the species? Our goals are hence to test the hypotheses of refugia along the oceanic–continental gradient and ‘oceanic’ adaptation, revealing how the species responded to past climate changes and provide information for collecting kiwifruit germplasm resources and conserve the species.

Results

Chloroplast DNA haplotype diversity

Three noncoding cpDNA regions were concatenated for individuals of *A. eriantha* and two outgroup species i.e. *A. fulvicoma* Hance and *A. chinensis* Planch., with a complete length of 1605 bp. Twenty-three cpDNA haplotypes were identified within *A. eriantha*, with eight being singletons (Additional file 1), including 16 substitutions and nine indels (1–19 bp). Our data showed intermediate levels of haplotype diversity ($h_T=0.498$) and nucleotide diversity ($\pi_T=9.1 \times 10^{-4}$) within the species. Significant phylogeographic structure ($N_{ST}=0.592 > G_{ST}=0.447$, $P < 0.001$) was detected. In the SAMOVA analyses, $K=3$ when F_{CT} values reaches the maximum (0.821), which means that three regional groups of populations were identified. These three groups are denoted as “Southeast edge”, “Southwest edge” and “Main part” according to their relative positions in the distribution area of the species. The “Southeast edge” and “Southwest edge” groups only contain a single population, i.e. HA in “Southeast edge” and LP in “Southwest edge”. Hierarchical AMOVA showed a great amount of variation (82.1%) occurred among the three regional population groups and only 6.6% presented differences among populations within groups, and 11.3% of the variation within populations (Additional file 2).

The network of 23 haplotypes displayed a star-like pattern, with eight haplotypes directly connected to H1 by one mutation (Fig. 1c). H1 was designated as an ancestral haplotype since it was most closely related to the outgroup and occupied an interior position in the network. The dominant haplotype H1 occurred in 156 individuals (70.6%) from 25 populations except for SQ, WH and HA,

three populations located in the Zhejiang-Fujian Hilly Region (Fig. 1a). However, most haplotypes (19/23) were restricted to a single population (Additional file 1). Two haplotype lineages were recognized based on the results of network analysis and Bayesian phylogenetic inference (Fig. 1b and c). The split between Lineage 1 and Lineage 2 was dated to 4.03 Ma (95% HPD: 2.47–5.57). Haplotype lineages showed distinct distribution ranges. The haplotypes in Lineage 2 (H8, H9, H10 and H11) were found exclusively in the Northwest part of the Zhejiang-Fujian Hilly Region. In Lineage 1, H20, H21 and H22 appeared only in the southern portion of the Zhejiang-Fujian Hilly Region, while H23 occurs in the Xuefeng Mountains.

The regional population group “Main part”, which contains all haplotypes in Lineage 2, was used for mismatch distribution analysis (MDA) and bayesian skyline plots (BSP) to estimate the possible population spatial expansion indicated by the clear star-like phylogeny of haplotypes in Lineage 2. The BSP showed a slight population expansion occurring between c. 3.5 and 1.5 Ma (Fig. 2a). However, this analysis cannot precisely estimate N_e because of the very broad confidence intervals. The result of the MDA showed that the observed distribution of pairwise differences among the haplotypes do not differ significantly from the expected distribution under the sudden-expansion model ($SSD=0.0007$, $P=0.6349$; $HRag=0.2707$, $P=0.5484$) (Fig. 2b). The time of spatial sudden expansion was estimated at 81,775 yr BP (95% CI: 10,112–95,404) based on the parameter τ (3.00; 95% CI: 0.371–3.5). The significantly negative values of Tajima’s D (-2.06 , $P < 0.05$) and Fu’s F_S (-16.32 , $P < 0.05$) also supported that the “Main part” had experienced regional expansions in the past. Since similar results were estimated for all populations of the species, we did not show them here. The small size and low number of haplotypes of the other regional groups “Southeast edge” and “Southwest edge” are unsuitable for estimating possible population spatial expansion, therefore no MDA, BSP and neutrality tests were conducted for them.

Six dispersal and six vicariance events were discovered by BBM (Bayesian Binary MCMC) analysis. The ‘east area’ was identified as the most likely ancestral area of the *A. eriantha* (Additional file 3).

Nuclear microsatellite loci data analysis

Among the 38 loci, two were detected as outliers by Arlequin (AET22 and AET167) and six by SamBada (UDK96-040, AET38, AET76, AET97, AET104 and AET167). Of these, locus AET167 was detected as an outlier by both tests and considered a locus under selection. The locus is significantly correlated with four bioclimatic variables: Temperature Seasonality (BIO4), Annual Precipitation (BIO12), Precipitation of Driest Month (BIO14) and

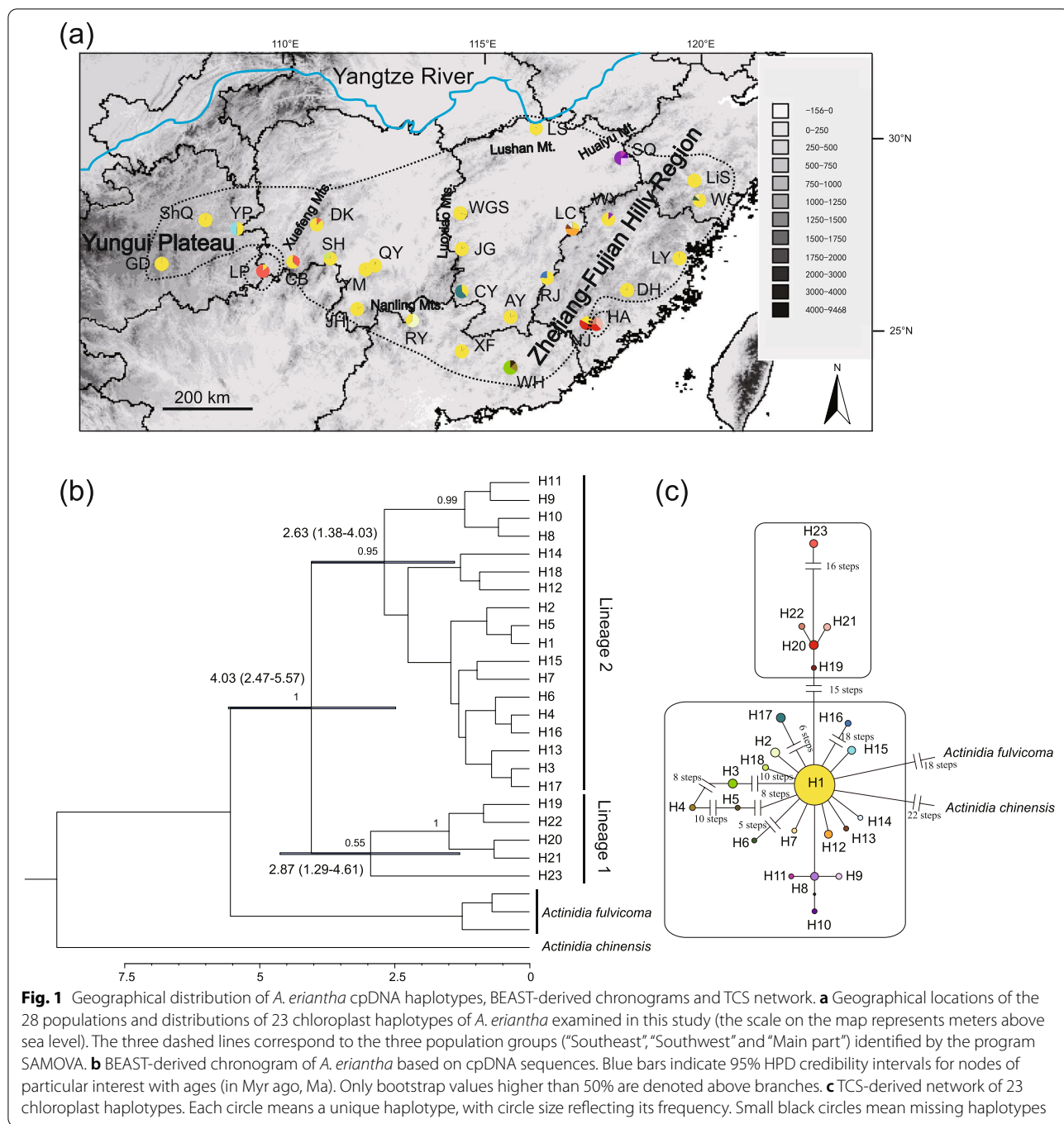


Fig. 1 Geographical distribution of *A. eriantha* cpDNA haplotypes, BEAST-derived chronograms and TCS network. **a** Geographical locations of the 28 populations and distributions of 23 chloroplast haplotypes of *A. eriantha* examined in this study (the scale on the map represents meters above sea level). The three dashed lines correspond to the three population groups ("Southeast", "Southwest" and "Main part") identified by the program SAMOVA. **b** BEAST-derived chronogram of *A. eriantha* based on cpDNA sequences. Blue bars indicate 95% HPD credibility intervals for nodes of particular interest with ages (in Myr ago, Ma). Only bootstrap values higher than 50% are denoted above branches. **c** TCS-derived network of 23 chloroplast haplotypes. Each circle means a unique haplotype, with circle size reflecting its frequency. Small black circles mean missing haplotypes

Precipitation of Coldest Quarter (BIO19). No genetic function was annotated for the sequence containing AET167 by GO analyses. In total, 31 loci (including 26 EST-SSRs and 5 genomic SSRs) displayed no evidence of selection in either outlier tests were therefore considered neutral loci.

We detected a relatively high total number of alleles per locus (mean 12.7, range 4-32). For all 38 loci, H_E ranged

from 0.293 to 0.938. Over all populations, 22 of the 38 loci displayed a significant heterozygosity deficit (Additional file 4). Significant deviations from Hardy-Weinberg equilibrium were detected in 86 of 1064 population-locus comparisons, yet no consistent patterns across loci or populations were evident. Null alleles were detected in 102 of 1064 combinations, which was above the expected number (53) by chance at the 5% level, albeit with no

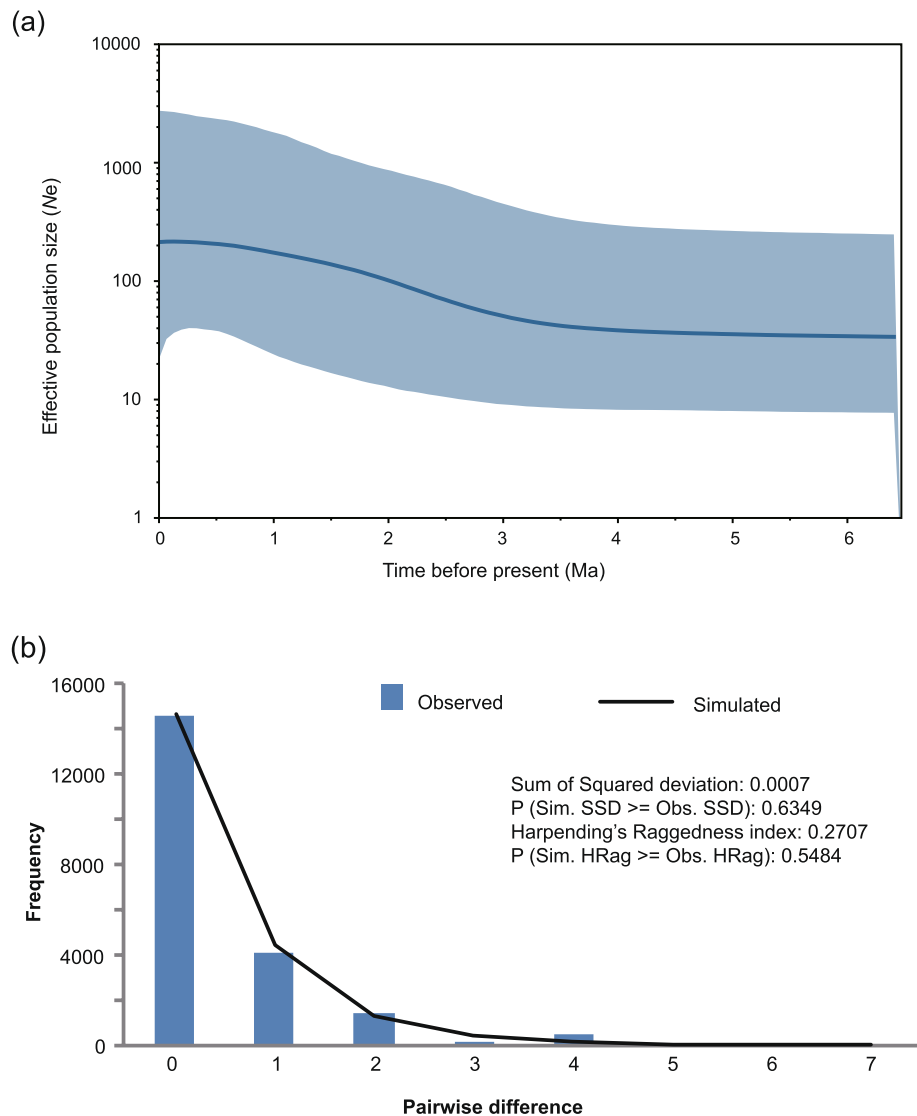


Fig. 2 The results of bayesian skyline plots (BSP) and mismatch distribution analysis (MDA) of the “Main part” inferred from *A. eriantha* cpDNA. **a** Bayesian skyline plots (BSP) estimated using BEAST2 v. 2.4. The thick solid blue line is the median estimate, and the area delimited by the light blue broadband represents the highest posterior density (HPD) 95% confidence intervals for N_e . **b** Mismatch distribution analysis (MDA) estimated in Arlequin v. 3.5

significant pattern specific to a population or locus. Linkage disequilibrium indicated significant deviations for 62 of 19,684 population-locus-locus combinations, which were lower than the expected number (984) by chance at the 5% level, suggesting no significant Linkage Disequilibrium among the 38 loci.

For each population across the 31 neutral loci, the genetic diversity parameters are listed in detail (Additional file 5). Values for A_e (mean 2.52, range 1.66-3.11), R_s (mean 2.781, range 2.014-3.240), H_o (mean 0.452, range 0.326-0.542) and U_{H_E} (mean 0.510, range 0.324-0.614) indicate moderate diversity within populations.

Significant heterozygosity deficit was displayed in 21 of the 28 populations. Higher diversity was identified in eastern *A. eriantha* populations (except for several marginal populations) (Fig. 3ab). Private alleles (A_p) occurred in all populations except YM and DK.

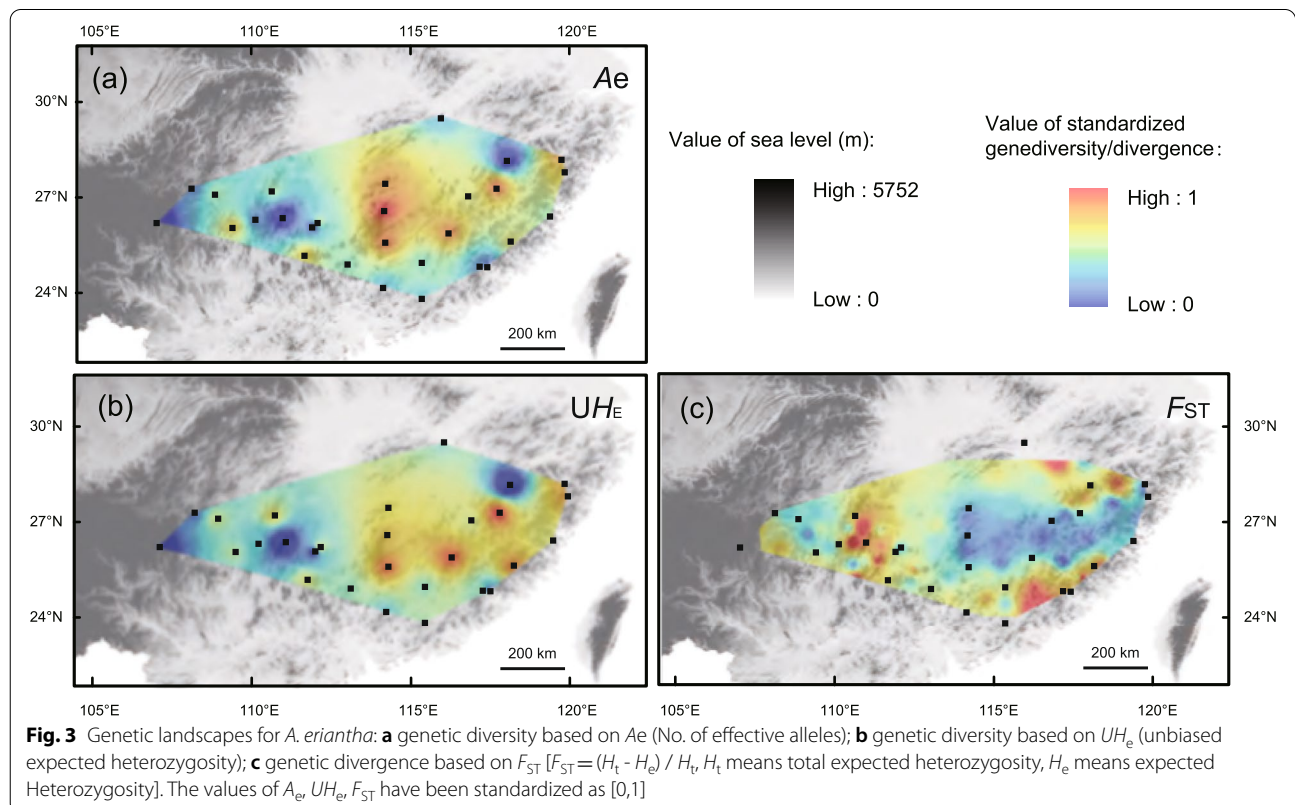
The genetic divergence pattern showed that eastern *A. eriantha* populations (except for several marginal populations) possessed lower values of genetic divergence than western populations (Fig. 3c). The Bayesian assignment indicated $K=2$ was the best when all populations were included (Additional file 6). Eastern populations (except for HA and NJ) were grouped into one cluster

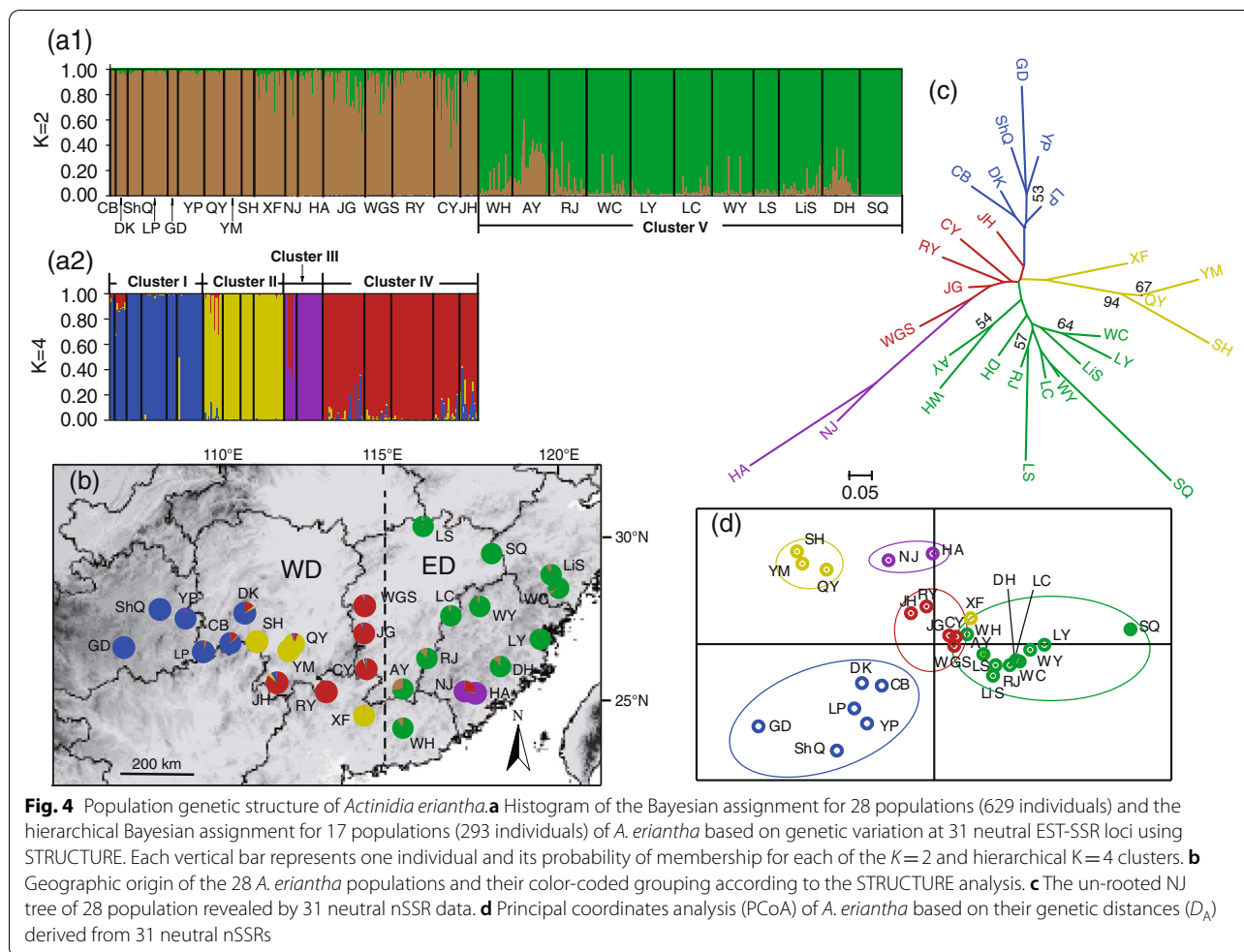
(Cluster V), and the rest were grouped into another cluster (Fig. 4a1). Subsequent hierarchical analyses split the rest of the species into four clusters: Cluster I, II, II and IV (Additional file 6 and Fig. 4a2). Overall, five clusters were identified, which were roughly consistent with the east-west distribution of the species (Fig. 4b). The PCoA and the unrooted NJ tree indicated a distinct differentiation among populations (Fig. 4c and d), in line with the findings of Bayesian assignment.

Historical gene flow among clusters (m_h) range from 0.001 (0-0.006) to 0.169 (0.155-0.173) (Additional file 7). Higher gene flow was estimated from cluster V to cluster I (0.082) and cluster V to cluster IV (0.169) unidirectionally. BayesAss yielded lower contemporary gene flow among most clusters than historical gene flow, with m_c ranging from 0.001 (0-0.003) to 0.047 (0.012-0.085) (Additional file 7).

Scenarios of historic processes were evaluated in a three-step analysis (see Additional file 8 for prior distributions). For the first set of analyses aiming to evaluate the relationship among the source genetic units (set A), the relative posterior probabilities calculated for each scenario provided the strongest statistical support for scenario A2 (0.5738, 95% CI: 0.5567-0.5908) (Additional file 9), suggesting that Cluster V diverged from Cluster III. Second, based on the results from set A, three

sets of scenarios were designed to analyze which source genetic unit Cluster I, II or IV diverged from (set B, C and D). The relative posterior probabilities calculated for each scenario of simulation set B, C and D, provided the strongest statistical support for Set B1 (0.6098, 95% CI: 0.5970-0.6227), C1 (0.5839, 95% CI: 0.5750-0.5927) and D1 (0.6344, 95% CI: 0.6156-0.6533) (Additional file 9), suggesting a common eastern origin (Cluster V) for the three colonization genetic units. Finally, Set E was calculated to test the relationship among colonization genetic units, based on the results from set B, C and D. Set E7 (0.2751, 95% CI: 0.2552-0.2950) was identified as the best-fit scenario (Fig. 5), indicated that Cluster I and Cluster IV diverged from Cluster V in chronological order, followed by the divergence of Cluster II from Cluster IV. We also estimated the divergence time and the population sizes for the five clusters (Additional file 10). The estimates in Set E were used since they are more credible when all populations were included in the set while the estimates from other sets may be biased due to missing segments of the populations. The time parameters are converted into years by multiplying generation time, which was set to 7 years for *A. eriantha*. Cluster V diverged from the ancestral population of *A. eriantha* Cluster III at approximately 460,000 generations ago (95% CI: 105,000 - 2,650,000) which corresponds





to c. 3.22 Ma (95% CI: 0.74 - 18.55). Then, Cluster I and Cluster IV diverged from Cluster V at approximately 294,000 (95% CI: 87,900 - 1,290,000) and 321,000 (95% CI: 107,000 - 785,000) generations ago, or c. 2.06 Ma (95% CI: 0.62 - 9.03) and 2.25 Ma (95% CI: 0.75 - 5.50), respectively. The split between Cluster II and Cluster IV was estimated to approximately 143,000 generations ago (95% CI: 38,000 - 360,000) or c. 1.00 Ma (95% CI: 0.27 - 2.52). No bottleneck was found by comparing the population sizes of N_i and these of N_{DBi} ($i = 1, 2, 4$). The population size of the species expanded along with divergence from 522,000 (N_3) to 3,145,000 ($N_1 + N_2 + N_3 + N_4 + N_5$).

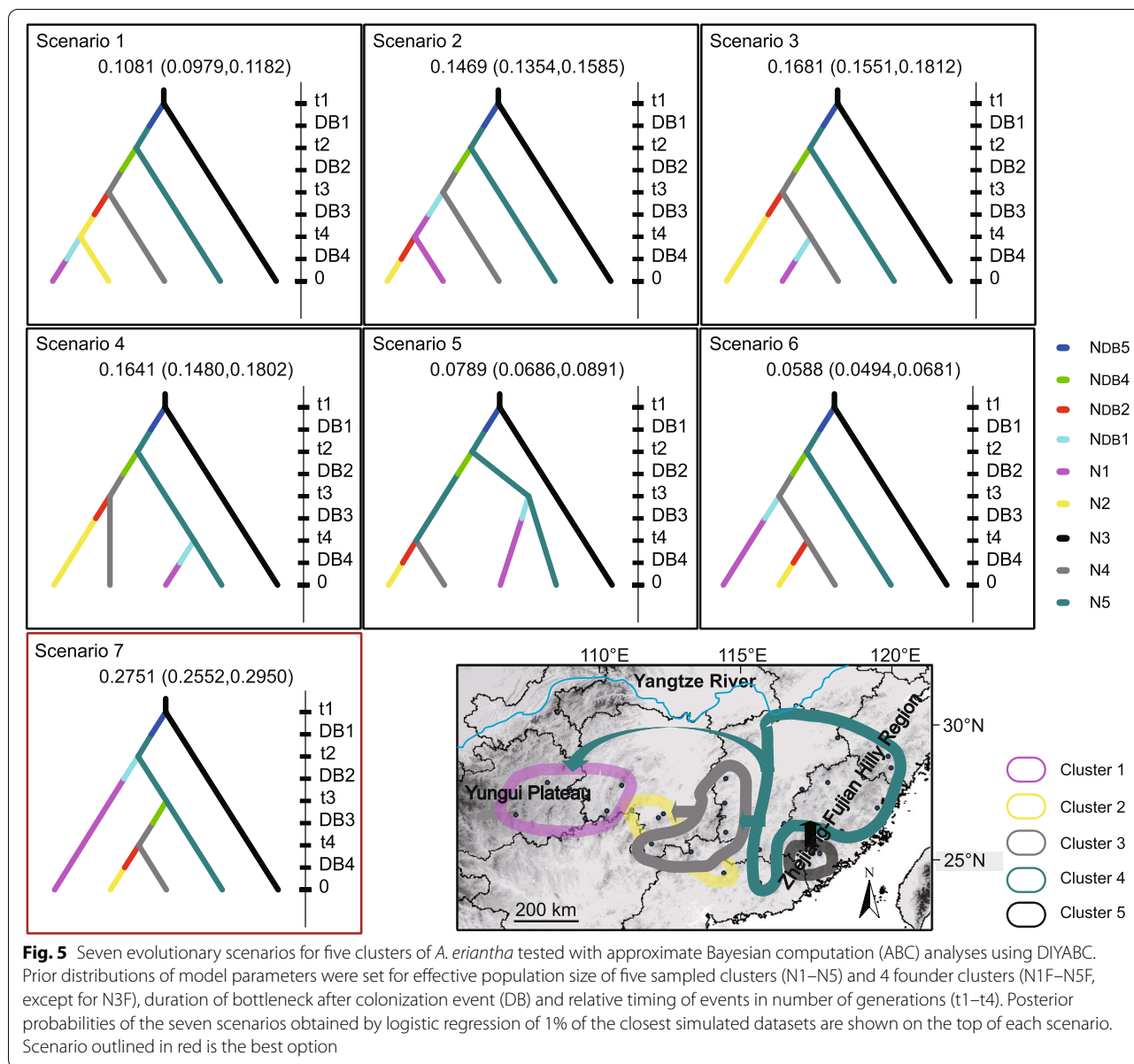
IBD and IBE analyses

The R^2 value for multiple matrix regression with randomization (MMRR) simulation based on the SSR neutral data was moderate ($R^2 = 0.127$, $P = 0.002$), suggesting a significant fit to this data, albeit not for cpDNA ($R^2 = 0.001$, $P = 0.968$). The MMRR results ($\beta_D = 0.279$, $P < 0.001$; $\beta_E = 0.139$, $P = 0.228$) indicated that geographic distance had a significant effect on neutral SSR

divergence, whereas the effects of environment factors were not significant (Table 1). For cpDNA divergence, there was no evidence of IBD or IBE ($\beta_D = 0.017$, $P = 0.823$; $\beta_E = 0.014$, $P = 0.933$; Table 1). The partial Mantel test also revealed similar results (SSR: $r_D = 0.267$, $P < 0.001$; $r_E = 0.137$, $P = 0.228$; cpDNA: $r_D = 0.016$, $P = 0.364$; $r_E = 0.013$, $P = 0.470$, Table 1).

Species distribution models and dispersal corridors

The Maxent models for *A. eriantha* had high predictive power and did not over-fit the present data (AUC values = 0.909). Our analysis indicated that Precipitation of Coldest Quarter (BIO19; importance of 44%), Temperature Seasonality (BIO4; 29%) and Mean Temperature of Wettest Quarter (BIO8; 25%) contributed the most to model development. The ensemble model for each period is illustrated in Fig. 6. The current potential distribution areas (defined as modelled suitability ≥ 0.6) of *A. eriantha* generally matches observed distributions, except for predicted but unsupported areas north of the Yangtze River. During the LIG, *A. eriantha* experienced a drastic



contraction compared to its current range with scattered areas in the Zhejiang-Fujian Hilly Region near the East China Sea and the eastern part of Taiwan province. Under LGM conditions, the estimated climatic suitability for *A. eriantha* underwent a radical range expansion with continuous suitable habitat from the Yungui Plateau to Taiwan. The future potential areas are contracted, fragmenting into several small suitable habitats scattered in the current distribution range of the species. The Zhejiang-Fujian Hilly Region (the current distribution area of Cluster V and Cluster III) near the East China Sea was always included in the predicted potential distribution areas of the species during the four glacial/interglacial

periods (Fig. 6). By comparing the isolines in the three climate layers during the four glacial/interglacial periods, we found that the expansion of *A. eriantha* was accompanied by higher precipitation during the coldest quarter, lower seasonal temperature difference and lower mean temperature during the wettest quarter throughout the distribution area (Additional file 11).

According to the western distribution (WD) and the eastern distribution (ED) range of the species, 158 occurrence points of *A. eriantha* were divided into ED (13 sampled locations and 63 herbaria specimens) and WD (15 sampled locations, 61 herbaria specimens and six location studied by Liu et al. [29] (see Additional file 12 for

Table 1 Results of Multiple Matrix Regression with Randomization (MMRR) analysis and partial Mantel test for SSR and cpDNA dataset of *Actinidia eriantha*

	MMRR						Partial Mantel test			
	Model		IBD		IBE		IBD		IBE	
	R^2	P	β_D	P	β_E	P	r_D	P	r_E	P
SSR	0.127	0.002	0.279	<0.001	0.139	0.228	0.267	<0.001	0.137	0.127
cpDNA	0.001	0.968	0.017	0.823	0.014	0.933	0.016	0.364	0.013	0.470

IBD Isolation by distance, IBE Isolation by environment

details). An additional 10 occurrences were not assigned into either group since they are located on the geographic boundary of the two distribution ranges. The Maxent models for both ED and WD had high predictive power and did not over-fit the present data (AUC values = 0.942

and 0.936, respectively). Precipitation of Coldest Quarter (BIO19; importance of 64.8%) and Temperature Seasonality (BIO4; 24.6%) contributed the most to model development for ED, while Temperature Seasonality (BIO4; 31.8%), Precipitation of Coldest Quarter (BIO19; 28.6%)

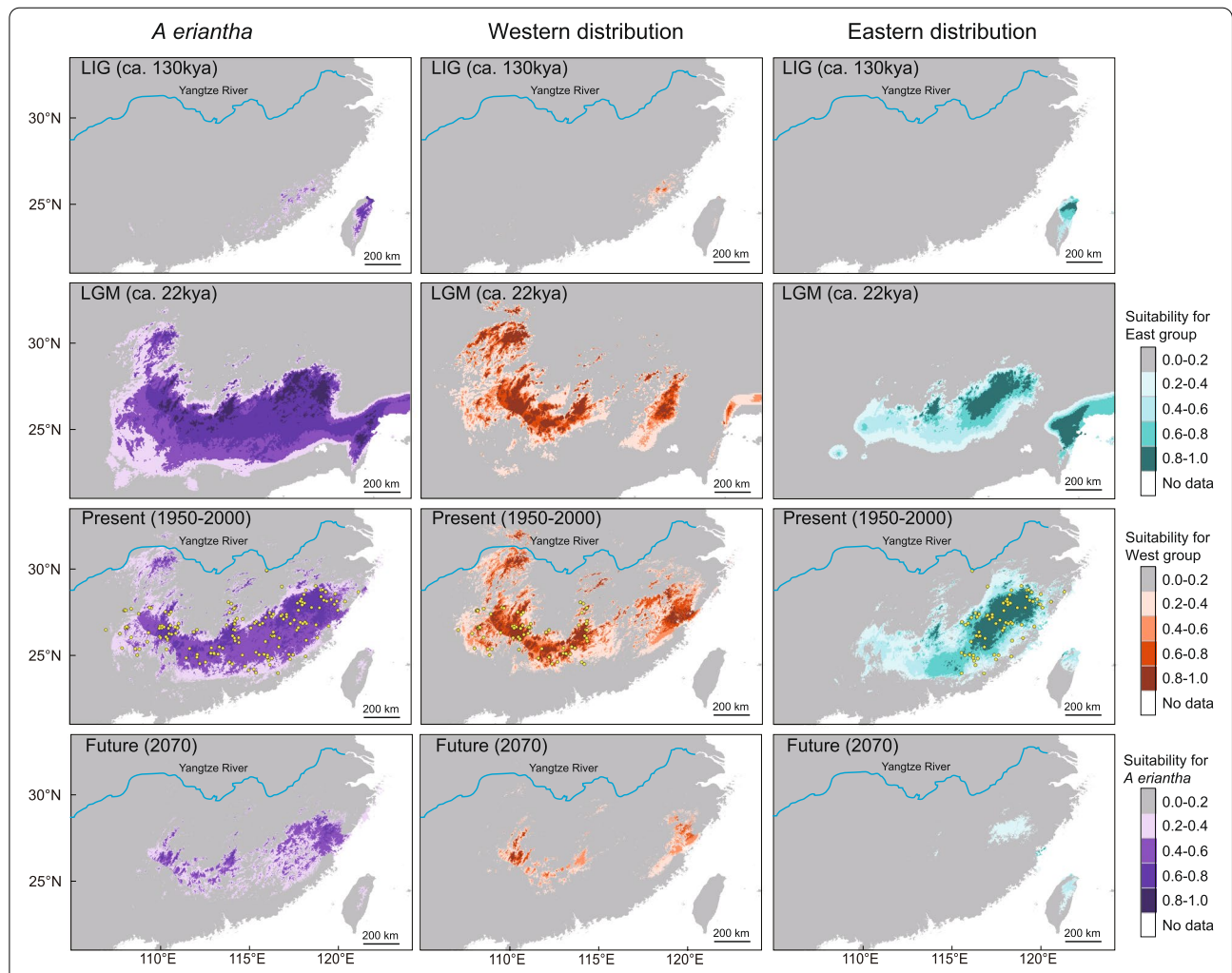


Fig. 6 Potential distributions of *A. eriantha*, West Distribution (WD) and East Distribution (ED) predicted using MaxEnt based on five bioclimatic variables representing the LIG, LGM, current and future climatic conditions, respectively. Warmer colors denote areas with a higher probability of presence. Dots show the extant occurrence points of the *A. eriantha*

and Mean Temperature of Wettest Quarter (BIO8; importance of 22.7%) contributed the most to model development for WD. The current potential distribution areas of WD not only match observed distributions, but also encompass part of the observed distribution of the ED and unsupported areas north of the Yangtze River. The current potential distribution areas of ED generally match observed distributions. Both WD and ED experienced a drastic contraction during the LIG, a radical range expansion during the LGM and a drastic contraction in the Future (2070). However, WD contracted into the Zhejiang-Fujian Hilly Region while ED contracted into Taiwan during the LIG. Under LGM conditions, potential distribution areas of both parts are similar to current ones, except for the areas of expansion to the northern part of Taiwan province. The future potential areas of WD are contracted to several small suitable habitats scattered in the current distribution range of the species, while the future potential areas of ED are contracted to several peninsulas and islands in Fujian province next to the East China Sea. Observed values of Schoener's D and Warren et al.'s I (0.44 and 0.72, respectively) were both significantly lower ($P < 0.01$) than expected from a random distribution (Additional file 13), suggesting the existence of niche differentiation between WD and ED.

The dispersal corridor result revealed a continuously east-west route during the LGM and the present but was hardly present during the LIG (Fig. 7).

Discussion

Genetic diversity and genetic structure of *Actinidia eriantha*

Patterns of genetic diversity of *A. eriantha* were investigated with neutral nuclear SSR loci and cpDNA data. At both the species and population levels, our results reveal a moderate level of microsatellite genetic diversity (mean $UH_E = 0.510$) and haplotype diversity ($h_T = 0.498$) across the 28 populations. The lower values of H_o than that of UH_E in most populations possibly resulted from inbreeding, which was confirmed by the significant heterozygosity deficit displayed in 21 of the 28 populations (Additional file 5). Genetic diversity values within populations of *A. eriantha* were slightly lower than that of other plants with a similar habitat, life history and breeding strategies [32]. However, genetic diversity at the population level of *A. eriantha* is likely underestimated. To our knowledge, EST-SSRs often display lower polymorphism than that of genomic SSRs [33]. In fact, genomic SSRs showed higher levels of genetic diversity than EST-SSRs in *A. eriantha* (Additional file 4) and 26 EST-SSRs were included to assess the genetic diversity of *A. eriantha* in the present study. The high genetic diversity

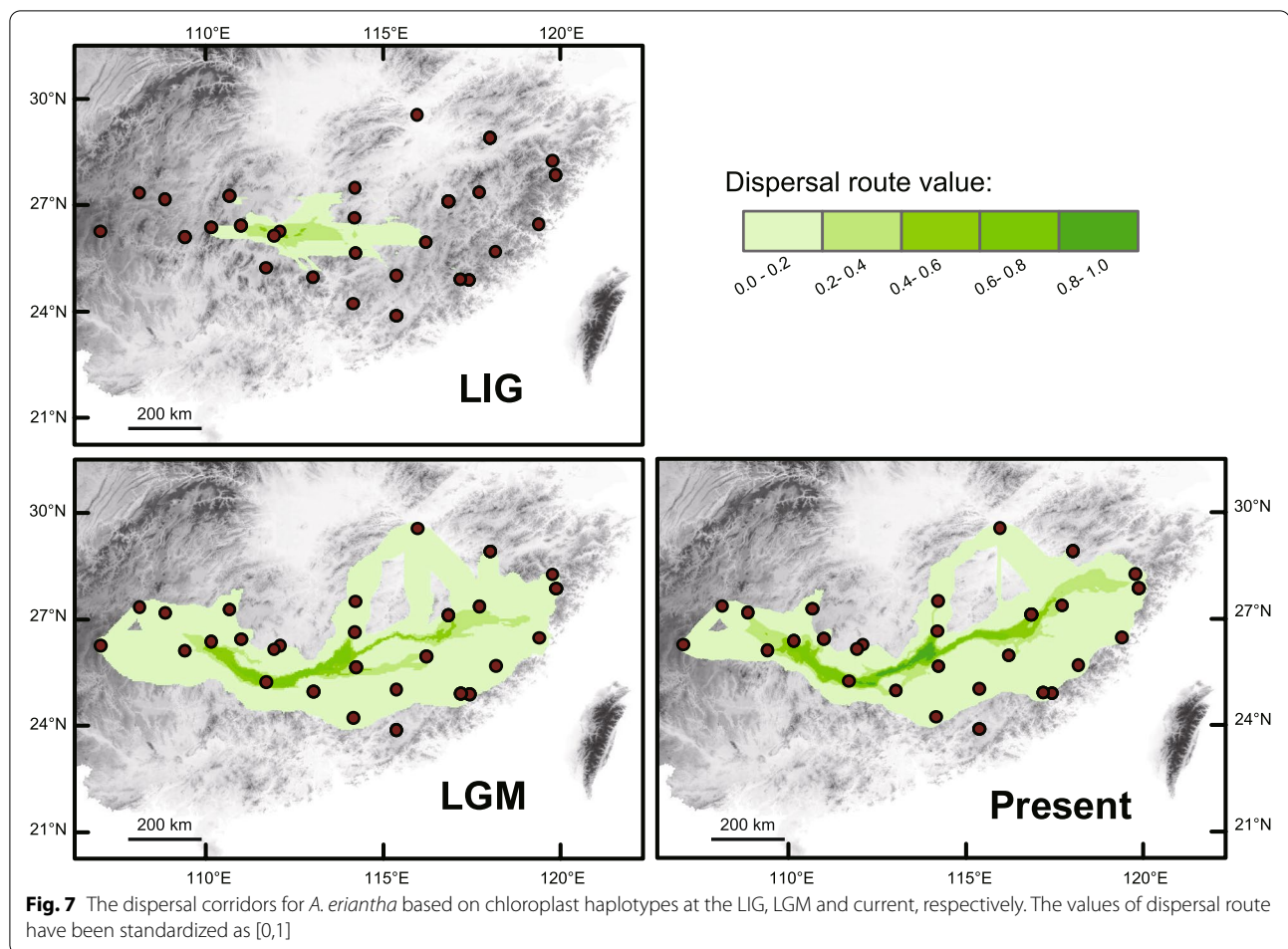
at the population level (mean $H_E = 0.763$) was previously observed in six populations located at the border between Hunan and Guangxi Province (nine genomic microsatellite loci) [29]. The high level of genetic diversity within populations for the predominantly insect-pollinated species is most likely explained by the outcrossing breeding system, high longevity and high inherent variability of the ancestral species of *A. eriantha*.

The genetic differentiation based on SSR markers ($F_{ST} = 0.177$) are almost identical to the mean values of genetic differentiation among populations of the late-successional ($F_{ST} = 0.17$) or long-lived ($F_{ST} = 0.19$) plants [32] and consistent with previous findings ($F_{ST} = 0.116$) [29]. The low levels of genetic differentiation may result from high levels of historical gene flow among the studied populations (Additional file 7). The significant correlation between genetic and geographic distances (Additional file 7) implies that gene flow is important in shaping the genetic differentiation. The fruit of *A. eriantha* is a relatively favorite desirable food source for frugivory animals. Also, seeds of *A. eriantha* can germinate readily upon maturity and are potentially capable of establishing new populations.

The significant phylogeographic structure was supported by chloroplast data according to SAMOVA (Fig. 1), Hierarchical AMOVA and the comparison values of N_{ST} and G_{ST} . Although three regional groups were identified by SAMOVA (Fig. 1) based on cpDNA, the two SAMOVA groups "Southeast edge" and "Southwest edge" probably represent areas of contact between haplogroups and may not reflect 'true' distinct biological units but mere artefacts due to low sample size. Thus, a closer inspection with nuclear DNA markers is needed. Five genetic clusters were revealed by the Bayesian assignment, PCoA and unrooted NJ tree analyses based on 31 neutral nuclear SSR loci (Fig. 4). Five clusters are roughly consistent with the east-west distribution of the species (Fig. 4b), demonstrating good geographic coherence. It is not uncommon that nuclear DNA markers can provide the high resolution required to explore the genetic structures of intra- or interspecies, for example in *Noccaea caerulea* (J. Presl & C. Presl) F. K. Mey [34] and *Betula* genus [35].

The refugium along the oceanic–continental gradient and the 'oceanic' adaptation

Both SDM and molecular phylogeography results revealed that the Zhejiang-Fujian Hilly Region next to the East China Sea is a refugium of *A. eriantha*. The Zhejiang-Fujian Hilly Region is always included in the predicted potential distribution areas of the species during the four glacial/interglacial periods in SDM (Fig. 6). In molecular analyses, the BBM results based on the



cpDNA indicated that *A. eriantha* spread from its eastern ancestral area to other distribution areas (Additional file 3) and the ABC analysis based on nuclear SSRs showed that Cluster I, II and IV in the WD of *A. eriantha* originated from Cluster V in the ED (Fig. 5). The high haplotype richness based on cpDNA (Fig. 1a, Additional file 1) and high genetic diversity based on neutral nuclear SSRs (Fig. 3, Additional file 5) in the Hilly Region are consistent with the expectation that repeated contraction and expansion to and from refugia leave genetic signals of high diversity in refugial areas and low diversity in areas of expansion [36, 37]. Higher levels of historical gene flow from Cluster V to Cluster I (0.082) and Cluster IV (0.169) unidirectionally also imply that the Zhejiang-Fujian Hilly Region is a refugium of *A. eriantha*.

The cpDNA results (Fig. 1 and Additional file 3) revealed that HA and NJ with Haplotypes H19-H22 in the southern part of the Zhejiang-Fujian Hilly Region is the ancestor of three populations LP, CB and DK with Haplotype H23 on/around the Xuefeng Mountains,

implying expansion from HA and NJ to the west distribution range. However, the haplotypes in the two populations are highly divergent from other haplotypes present in the Zhejiang-Fujian Hilly Region refugium. The localised highly divergent populations (Fig. 1b) indicate persistent isolation with other populations in the refugium from middle Pleistocene with no further expansion [38]. Thereby the populations have not been the source of major glacial recolonization as demonstrated by the ABC analysis based on neutral nSSRs (Fig. 5 and Additional file 9). The phylogeographic lineages remain geographically distinct within the refugium, which fit the 'refugia within refugia' pattern that has been described in the Iberian [39]. Such pattern revealed the internal complexity of the Zhejiang-Fujian Hilly Region as a refugium along the oceanic–continental gradient.

The preference for higher humidity during the coldest quarter and lower seasonal temperature difference of *A. eriantha* and its dispersal along the longitude imply a refugium along the oceanic–continental gradient and

'oceanic' adaptation of the species. SDM analysis indicated that Precipitation of Coldest Quarter (BIO19; importance of 44%), Temperature Seasonality (BIO4; 29%) and Mean Temperature of Wettest Quarter (BIO8; 25%) contributed the most to model development. By checking the isolines in the three climate layers during the four glacial/interglacial periods, we found that *A. eriantha* prefers higher precipitation during the coldest quarter, lower seasonal temperature difference and lower mean temperature during the wettest quarter (Additional file 11). The precipitation during the coldest quarter is similar to that of the present (Additional file 11), despite the climate of this region during the LGM (last glacial maxima) being dryer by c. 400–600 mm/yr [18, 19]. According to the location and type of the refugium of *A. eriantha*, we inferred that some or at least one of the populations of *Actinidia* fragmented in the Zhejiang-Fujian Hilly Region adapted 'oceanic' climate and formed *A. eriantha* during late-Miocene/early-Pliocene. Although a dipole-type circulation pattern of atmosphere during winter caused by the Tibetan Plateau and a lag time in spring warming between land and sea during February generate the moist environment in the coldest season in the east subtropical China [40–42], the moist level should always decrease from the coast to inland areas during glacial cycles as ranges of mountains in the Zhejiang-Fujian Hilly Region intercept clouds and cause major local increase in rain and snowfall. Then the species expand from the refugium repeatedly with the higher moist during coldest quarter and lower seasonal temperature difference during wettest quarter during the glacial periods.

The biogeographic history of *Actinidia eriantha*

Actinidia eriantha in the refugium dispersed westward along a continuously east-west dispersal route from the Zhejiang-Fujian Hilly Region next to the East China Sea (Additional file 3, Figs. 5, 6 and 7). The corridor role of the Wuyi, Nanling, and Luoxiao Mountains during the late Quaternary has been previously mentioned [8, 13].

Actinidia eriantha went through repeated glacial expansions. The BBM analysis based on cpDNA estimated six dispersal and six subsequent vicariance events in the species. The BSP, MDA and neutrality tests based on cpDNA confirmed the indication of possible expansions in Lineage 2 by the star-shaped haplotype TCS network based on cpDNA (Fig. 1). The BSP showed that a slight population expansion occurred between c. 3.5 and 1.5 Ma (Fig. 2a). Based on neutral nuclear SSRs, DIYABC revealed that expansions along with the divergence of the five genetic clusters happened between 3.22 and 1.00 Ma, which is close to the expansion period estimated in BSP based on cpDNA. The two genetic data sets

used here provide signals of demographic events which have occurred at the same time scales. The result of MDA showed that populations went through a sudden expansion 81,775 yr BP (95% CI: 10,112–95,404), which coincides with the early part of the last glacial period over the past 800,000 years (c. 115,000 – c. 11,700 years ago) [43]. Additionally, *A. eriantha* underwent a significant range expansion during the LGM (21 kya BP) based on SDM. Accumulating evidence suggests glacial expansion was not a rare event in subtropical China (e.g. [12, 44]).

The more highly fragmented distribution area of the species is evident from the LGM to the future based on SDM (Fig. 6). This is supported by lower contemporary gene flow among most clusters than historical gene flow (Additional file 7) suggesting that gene flow among these clusters is now more restricted.

Important roles of geographic distance and environment distance for intraspecific divergence

Both geographic distance and environment distance have important roles for intraspecific divergence of *A. eriantha*. Environment effects can contribute to patterns of genetic structure of species either genome-wide or only in particular genomic regions [45]. Environment factors played an important role in adaptive differentiation of *A. eriantha* since the result of the identity test supports the existence of niche differentiation between the two groups (Fig. 6 and Additional file 13). Contribution analysis of SDM revealed both precipitation and temperature made great contributions to model development. In addition, the divergence times of the two lineages detected in *A. eriantha* based on three cpDNA fragments were dated to 4.03 Ma (95% HPD: 2.47–5.57 Ma) (Fig. 1b), which coincide with a turning point of drastic climate fluctuation during the middle Pliocene [46]. However, environment distance has no significant correlation with genetic differentiation of *A. eriantha* based on neutral loci (Table 1), suggesting that the effects of environment factors have not yet spread genome-wide. The predicted patterns for IBD are that genetic differentiation at neutral loci increases with increasing geographic distance, as a consequence of reduced gene flow as geographic distances increase [47]. The pattern has been proved by the IBD test ($\beta_D = 0.279$, $P < 0.001$) (Table 1).

Conclusions

In summary, this is the first integrated evidence of a refugium of *Actinidia eriantha* along the oceanic–continental gradient and 'oceanic' adaptation of the species in the subtropical China. Molecular phylogeography and species distribution modelling revealed west-east geographically distinct divergence of *Actinidia eriantha*. Both geographic distance and climate difference all have

important roles for intraspecific divergence of the species. The Zhejiang-Fujian Hilly Region was demonstrated to be a refugium along the oceanic–continental gradient and fit the ‘refugia in refugia’ pattern. After originating from the Zhejiang-Fujian Hilly Region with ‘oceanic’ adaptation, *A. eriantha* expanded to its western distribution range along the dispersal corridor repeatedly during the glacial periods. This study gives a deeper understanding of the refugia in subtropical China and will contribute to the conservation and utilization of kiwifruit wild resources since the identification of the refugium and the genetic diversity it harbour enables the refugium to be targeted for protection and collection in the context of climate change.

Materials and methods

Sample collection and DNA extraction

The leaves of 629 individuals from 28 locations were sampled across the entire distribution range of *A. eriantha* (Table 2, Fig. 1). Nine to 35 individuals in each population were sampled, except for populations CB and GD on the periphery of the range, in which only five and eight individuals were found, respectively. Three individuals of *A. fulvicoma* and one individual of *A. chinensis* were used as outgroups. Fresh leaves of outgroup species were collected from the Wuhan Botanical Garden, CAS. Voucher specimens representative of all samples are stored at the Herbarium of Wuhan Botanical Garden. Total genomic DNA was extracted from silica-dried leaves using a modified CTAB method [48]. Quality and concentration of the DNA were confirmed using 1% agarose gel electrophoresis and NanoDrop 8000 (Thermo Fisher Scientific, Waltham, MA, USA).

Laboratory protocols

For the phylogeographic cpDNA analysis, we sequenced three non-coding intergenic spacer (IGS) regions: *ndhF-rpl132* (F: ACAGGAACTGGAAGTGGAACAA; R: TTG GTCAAGGTCGAGAAAGAA), *rps16-trnQ* (F: GTC GCACGTTGCTTTCTACC; R: TAGCTGCGTTGT CCGAATCT), and *trnE-trnT* (F: ATACTTGCCCGA CCGACATC; R: GAACCGACGACTTACGCCTT). The cpDNA of *A. eriantha* is inherited paternally [49]. In total 221 individuals from 28 populations of *A. eriantha*, three individuals of *A. fulvicoma* and one of *A. chinensis* were sequenced. Polymerase Chain Reactions were conducted in 20 µl reactions containing 10 µl 2× *Taq* PCR Master-Mix (Biotek), 0.5 µl each primer (0.2 µM), 1 µl template DNA (ca. 50–100 ng) and 8 µl ddH₂O. Thermo-cycling conditions were as follows: 94 °C for 4 min; 35 cycles of 94 °C for 30 s, 58 °C for 60 s, and 72 °C for 60 s; a final 10 min extension at 72 °C. Sequences were generated with an ABI 377XL DNA sequencer (Applied Biosystems).

For nuclear DNA analysis, 629 individuals were genotyped at 38 nuclear microsatellite loci (Additional file 4) with 32 EST-SSRs described by Guo et al. [31] and six nuclear SSRs screened from *A. chinensis* (five genomic SSRs: UDK96-030, UDK96-026, UDK96-040, 751, 761 [50, 51] and one EST-SSRs [52]). Amplification of SSR loci followed the protocol in Guo et al. [31]. Fluorescently labelled PCR products were spiked with the internal size standard GeneScan 500 LIZ and separated on a 3730xl DNA Analyzer (Applied Bio systems). Alleles were scored manually in GeneMarker v. 2.2 (SoftGenetics, Pennsylvania, USA). Microsatellite quality was checked using MSAnalysr [53].

Data analysis

Analysis of cpDNA variation

All sequences were checked using Finch TV v. 1.4 (<https://digitalworldbiology.com/FinchTV>). Three cpDNA regions were aligned and trimmed separately in MEGA v. 6 [54] and combined using FaBox v. 1.5 (<https://users-birc.au.dk/~palle/php/fabox/index.php>). Haplotype (*h*) and nucleotide (π) diversity were calculated for each population and the overall species using DnaSP v. 5 [55]. Two parameters for population differentiation (G_{ST} , N_{ST}) were analyzed using the program PermutCpSSR v. 1.2 (<http://www.mybiosoftware.com/tag/permutcpsr>).

To infer possible regional groups of populations, spatial analysis of molecular variance (SAMOVA) of cpDNA was implemented in SAMOVA v. 1.0 [56]. The most likely number of groups (*K*) was determined by repeatedly running the program SAMOVA with 2–28 groups and choosing those partitions with a maximum value of differences among groups of populations (F_{CT}). Hierarchical analysis of molecular variance (AMOVA) [57] was also conducted in Arlequin v. 3.5 [58] to quantify the proportion of total genetic variance explained by differences between regional population groups (as identified by SAMOVA) and between populations within groups. Significance of variance components was tested with 10,000 permutations.

Genealogical relationships of identified haplotypes were inferred from a 95% statistical parsimony network constructed in TCS v.1.2 [59]. To identify cpDNA lineages and estimate divergence time among lineages, we performed Bayesian phylogenetic inference on the haplotypes with a relaxed clock model in BEAST2 v. 2.4 [60]. A Yule prior was applied for the inter-species relationships and a coalescent prior assuming constant population size for the intra-species relationships. The GTR model was selected as the best-fit substitution model of molecular evolution using jModelTest v. 2.1 [61]. The diversification in *Actinidia* has been previously elucidated [62].

Table 2 Characteristics of sampled populations of *Actinidia eriantha*

Population code	Location	Longitude (E)	Latitude (N)	Altitude (m)	Sample size
WH	Wuhua County, Guangdong Prov.	115°23'	23°52'	686	28
LY	Luoyuan County, Fujian Prov.	119°24'	26°27'	529	35
DH	Dehua County, Fujian Prov.	118°11'	25°40'	1001	30
JG	Mount Jinggang, Jiangxi Prov.	114°11'	26°36'	1000	33
LiS	Lishui city, Zhejiang Prov.	119°46'	28°15'	365	33
SQ	Mount Sanqing, Jiangxi Prov.	118°03'	28°12'	589	34
AY	Anyuan County, Jiangxi Prov.	115°23'	25°00'	433	29
RY	Ruyuan County, Guangdong Prov.	113°03'	24°57'	820	33
RJ	Ruijin City, Jiangxi Prov.	116°13'	25°56'	400	30
NJ	Nanjing County, Fujian Prov.	117°12'	24°53'	695	10
WC	Wencheng County, Zhejiang Prov. ProvProvince	119°52'	27°30'	400	35
XF	Xinfeng County, Guangdong Prov.	113°03'	24°57'	594	25
CY	Chongyi County, Jiangxi Prov.	114°14'	25°38'	478	21
YM	Mount Yangming, Hunan Prov.	111°56'	26°7'	1179	15
SH	Mount Shunhuang, Hunan Prov.	111°00'	26°24'	668	9
JH	Jianghua County, Hunan Prov.	111°42'	25°13'	420	15
LS	Mount Lu, Jiangxi Prov.	115°58'	29°33'	1080	20
WY	Mount Wuyi, Jiangxi Prov.	117°43'	27°20'	500	33
LC	Lichuan County, Jiangxi Prov.	116°50'	27°05'	277	30
ShQ	Shiqian County, Guizhou Prov.	108°08'	27°20'	977	12
LP	Liping County, Guizhou Prov.	109°25'	26°06'	450	20
CB	Chengbu County, Hunan Prov.	110°09'	26°22'	1485	5
DK	Dongkou County, Hunan Prov.	110°40'	27°14'	535	9
GD	Guiding County, Guizhou Prov.	107°03'	26°15'	1113	8
YP	Yuping County, Guizhou Prov.	108°52'	27°09'	506	21
QY	Qiyang County, Hunan Prov.	112°06'	26°15'	146	15
WGS	Mount Wugong, Jiangxi Prov.	114°13'	27°29'	619	22
HA	Hua'an County, Fujian Prov.	117°26'	24°52'	861	20

Therefore, we can take advantage of the timing of kiwi-fruit diversification as calibration points when doing our dating analyses. Two calibration points were used with a normal distribution prior: the split between *A. chinensis* and the two other species of 11 Myr (95%CI: 4.7-17.3 Myr) and the split between *A. eriantha* and *A. fulvicoma* of 5.5 Myr (95%CI: 4.2-6.8 Myr) [62]. We did not use the calibration point of the split between *A. eriantha* and its closest relative *A. latifolia* since *A. latifolia* shows up in the ingroup of *A. eriantha* in the tree constructed with the three cpDNA markers. The length of the Markov Chain Monte Carlo (MCMC) algorithm was set to 1 billion generations, sampling every 100,000 generations with the first 20% discarded as burn-in. We checked the convergence of parameters using Tracer v. 1.6 [63]. The consensus tree was analyzed using TreeAnnotator and visualized in FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

We used bayesian skyline plots (BSP) in BEAST2 v. 2.4 and the mismatch distribution analysis (MDA) in Arlequin v. 3.5 to detect population spatial expansion events. BSP was conducted with the same settings as in previous analyses except that the priors setting was changed to Coalescent Bayesian Skyline. BSP has the advantage of not assuming any demographic scenario a priori [64]. The goodness-of-fit under a sudden-expansion model was tested with the sum of squared deviations (SSD) and Harpending's raggedness index (HRag) [65] for MDA. The MDA-derived spatial expansion parameter (τ) was converted into generation time (T) using the following equation: $T = \tau/2\mu$ [66], where μ is the neutral mutation rate of the entire cpDNA sequence per generation. The value for μ was calculated as $\mu = uk$, where u is the substitution rate (here, 1.0×10^{-8} substitutions/site/generation(s/s/g) [62], and k represents the sequence length of the cpDNA region (here, 1605 bp). Finally, the expansion time was calculated assuming a generation

time of 7 years for *A. eriantha* under natural conditions [62]. We also calculated Tajima's D [67] and Fu's F_S [68] to assess possible expansions using Arlequin v. 3.5.

The ancestral areas were reconstructed using the BBM (Bayesian Binary MCMC) implemented in RASP v. 3.0 [69]. In total, 10,000 BEAST-generated trees and one consensus tree were used as input with outgroups removed and a 10% burn-in. Three regions (west, middle and east areas) were defined according to the precision of consensus tree and the distribution range of the species (Additional file 3). The F81 model (variable base frequencies, all substitutions equally likely) [70] was used according to Akaike Information Criterion (AIC: F81 = 4949, JC = 5335) estimated using jModelTest.

Nuclear data analysis

To identify putative locally adaptive loci that may affect analysis of neutral population structure and demographic history, we used two methods: genetic-environment association analysis in SamBada v. 0.5.3 [71] and differentiation outlier in Arlequin with a significance threshold of 0.01 corrected using the Bonferroni method. These two methods have recently become widely used to test for loci under selection [72]. A total of 19 bioclimatic variables (<http://www.bioclim.org>) served as environmental variables. Two sets of loci were obtained: loci under selection and neutral loci. A locus was considered to be under selection if both outlier tests showed significance. The correlation between the outlier loci and environment factors was checked using SamBada. Gene Ontology (GO) annotations [73] of the sequences containing the outliers were conducted. A locus was considered to be neutral if both tests showed no significance. Successive nuclear data analyses were based only on neutral loci except for the analysis of characterization of each locus.

Polymorphism of each microsatellite locus, i.e. the number of observed alleles (N_A), the observed (H_O) and the expected heterozygosity (H_E), and polymorphism information content (PIC) was calculated using CER-VUS v. 3.0 [74]. The mean fixation coefficient F_{IS} was calculated using FSTAT v. 2.9 [75]. Significant deviation from Hardy-Weinberg equilibrium (HWE) induced by heterozygote deficiency was tested across all populations and each population-locus combination using 10,000 permutations in GENEPOP v. 4.2 (<http://genepop.curtin.edu.au/>). Significance was corrected for multiple testing using the Bonferroni method. To account for any departure from HWE due to the presence of null alleles in each locus, we estimated their frequency within each population in FreeNA with a cutoff value of 0.2 [76]. Linkage disequilibrium among the 38 loci was evaluated in GENEPOP for each population-locus-locus combination using 10,000 permutations.

Intrapopulation genetic variation was estimated by the total number of alleles (A_T), the number of private alleles (A_p), the effective number of alleles (A_e), the observed (H_O) and the expected heterozygosity (H_E), unbiased expected heterozygosity (UH_E) [77], and the inbreeding coefficient F_{IS} across 31 neutral loci using GeneAlex v. 6.5 [78]. The allele richness (R_S , standardized for four individuals using rarefaction) was calculated using FSTAT. The two genetic parameters UH_E and R_S were used for removing the bias caused by the strong variation in the sample size. Genetic diversity based on standardized A_e and UH_E were visualized as genetic landscapes in ArcGIS v. 10.2 (Environmental Systems Research Institute, Inc., Redlands, CA), using the Genetic Landscapes Toolbox [79]. Significant deviations of Hardy-Weinberg equilibrium induced by heterozygote deficiency were tested in GENEPOP for each population across 31 neutral loci using 10,000 permutations. Pairwise population F_{ST} was estimated using GeneAlex. Genetic distance based on standardized F_{ST} was visualized as divergence landscape.

Several methods were used to investigate the genetic structure of *A. eriantha*. A principal coordinate analysis (PCoA) of populations was performed in GenAlex and the first two PC axes (PC1 and PC2) were used to explore the genetic relationships among populations. The genetic relationships were also evaluated by generating a neighbor-joining (NJ) network based on D_A distances [80] using POPTREE v. 2 [81]. The most likely number of clusters was inferred by Bayesian clustering implemented in STRUCTURE v. 2.3 [82]. Assuming admixture and independent allele frequencies, we set up a run with a burn-in period of 100,000 iterations followed by 500,000 recorded iterations for $K=1$ to $K=28$ clusters and 10 runs per K value. The most probable number of clusters was determined using the ΔK approach [83] in STRUCTURE HARVESTER [84].

Migrate-n v. 3.6 [85] was used to estimate levels of historical gene flow between each pair of clusters by calculating the mutation-scaled effective immigration rate (M) with the Brownian motion approximation. Three independent analyses were run, recording every 100 steps with 500,000 genealogies and a 10,000 genealogy burn-in. Immigration rate (m) was calculated as $m = M\mu$ with μ being the mutation rate (estimated for nSSRs as 3×10^{-4} , [86]). BayesAss v. 3.0 [87] was used to estimate the contemporary counterpart by reckoning each pairwise m . A total of 5,000,000 MCMC iterations were run with 20% burn-in.

We used ABC simulations in DIYABC v2.0 [88] to determine the historic process involved in the settlement of clusters identified by STRUCTURE based on the 31 neutral SSR loci. Due to computational limitations and

the infinite number of possible scenarios when numerous populations are considered, inferences were based on finite set of genetically and geographically delimited groups. According to the phylogeography results using of cpDNA, patterns of genetic divergence revealed by SSR, and SDM during past, present and future, we denoted clusters located in the eastern distribution (ED) range (Cluster V and Cluster III) as source genetic units and evaluated the relationship among the source genetic units. Furthermore, we detected the origin of each colonization genetic unit located in the western distribution (WD) range, and the order of divergence of the units. Scenarios were evaluated in a three-step analysis. First, a set of scenarios aimed to analyze the relationship among the source genetic units (set A). Second, based on the results from set A, three sets of scenarios were designed to analyze which source genetic unit Cluster I, II or IV diverged from (set B, C and D). Finally, Set E was calculated to test the relationship among colonization genetic units and their source genetic units, based on the results yielded by set B, C and D. A list of all parameters and prior distributions used to model scenarios is summarized in Additional file 8. The type I and mean type II error for the most likely scenarios were calculated to estimate the statistical power [89]. The time parameters were estimated in generations. As colonization may generate strong demographic bottleneck leading reduction in genetic diversity, times of demographic bottleneck were considered in our set of scenarios.

Isolation by distance and isolation by environment analyses based on both neutral SSR loci and cpDNA dataset

To evaluate the effect of geographic and environmental conditions on genetic divergence, we tested for isolation by distance (IBD) and isolation by environment (IBE) based on neutral SSR loci and cpDNA dataset. IBE based on potentially adaptive loci was not tested since only one outlier locus was obtained. Two methods were used: multiple matrix regression with randomization (MMRR) [90] using the R function 'MMRR' and partial Mantel test in the package *vegan* v. 2.4 [91]. We used pairwise $F_{ST}/(1-F_{ST})$ [92] to represent population pairwise genetic distances. Geographic distances were shown by the log₁₀ of geographic distances between pairs of populations. For environmental variables, we obtained 19 bioclimatic variables (www.worldclim.org). To reduce bioclimatic covariance, principal component analysis (PCA) was conducted using the R package *vegan*. The first three PC axes (PC1, PC2 & PC3) explaining >89% of the total variation were used to compute Climatic (Euclidian) distance.

Species distribution models and dispersal corridors analyses

To infer the potential distribution areas of *A. eriantha* across a complete glacial-interglacial cycle, 168 occurrence points of *A. eriantha* including 28 sampled locations, six location studied by Liu et al. [29] and 134 herbaria specimens (Additional file 12) from the Chinese Virtual Herbarium (<http://www.cvh.ac.cn/>) were used to compute species distribution models (SDM) in Maxent v. 3.2 [93] based on Future (2080), Current (1950-2000), LGM (21 kya BP), and LIG (120-140 kya BP) climatic maps (<http://www.worldclim.org>). Model performance was evaluated using the area under the Receiver Operating Characteristic Curve (AUC). To avoid over-fitting, we selected five bioclimatic variables ($|r| > 0.8$): BIO2 (Mean Diurnal Range), BIO4 (Temperature Seasonality), BIO8 (Mean Temperature of Wettest Quarter), BIO16 (Precipitation of Wettest Quarter) and BIO19 (Precipitation of Coldest Quarter), based on correlation analysis using ENMTools v. 1.3 [94] and their contributions to SDM. Relative contributions of the environmental variables to model development were estimated in Maxent. Model predictions were visualized in DIVA-GIS (<http://www.diva-gis.org/gdata>). To determine whether the WD and ED part of the species occupy identical climatic environments ('niches'), we performed SDM in Maxent and niche identity tests in ENMTools. A one-tailed t -test was used to estimate the significant difference between observed values of Warren et al.'s I [95] and Schoener's D [96] and 100 randomized distributions of D and I generated by randomly assigning samples to either part.

Dispersal corridors at the current time, LGM, and LIG were mapped by applying the CLCP method and SDM toolbox [97] in ArcGIS. To create the ecological dispersal network, we converted SDM generated above to dispersal cost layers. Chloroplast haplotype pairwise population networks were generated by taking the sum of the least cost paths (LCPs) among all shared and sister haplotypes from different localities in ArcGIS. Standardized dispersal corridor layers of *A. eriantha* were established using SDM toolbox.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-022-03464-5>.

Additional file 1. Genetic characteristics of cpDNA in 28 *Actinidia eriantha* populations.

Additional file 2. Analysis of molecular variance (AMOVA) for 28 sampled populations of *Actinidia eriantha* based on cpDNA.

Additional file 3. The ancestral areas were reconstructed using the BBM (Bayesian Binary MCMC) method. (a) The ancestral areas were reconstructed using the BBM implemented in RASP v. 3.0. Three regions

(west, middle and east areas) were defined according to the precision of consensus tree and the distribution range of the species. (b) Geographical locations of the three regions and the most likely dispersal direction.

Additional file 4. Characteristics of 38 nuclear microsatellite loci.

Additional file 5. Genetic diversity of each population based on 31 neutral nuclear microsatellite loci.

Additional file 6. The most probable number of clusters determined using the Delta K approach. (a) When $K = 2$ for all individuals, Delta K has the highest value. (b) When $K = 4$ for western subset (plus HA and NJ) of individuals, Delta K has the highest value.

Additional file 7. Migration rates (m) across the five clusters of *Actinidia eriantha* based on 31 neutral nuclear SSRs.

Additional file 8. Prior distributions of the parameters used in DIYABC.

Additional file 9. Set A, B, C and D of scenarios in DIYABC of *Actinidia eriantha*. Posterior probability of each scenario obtained by logistic regression of 1% of the closest simulated datasets is shown on the top of the scenario. Scenario outlined in red is the best option.

Additional file 10. Estimations of posterior distributions of parameters revealed by DIY-ABC for the best scenarios of set A, B, C, D and E, respectively.

Additional file 11. The isolines in the three climate layers Bio 4, Bio 8 and Bio 19.

Additional file 12. Information of 134 herbaria specimens of *Actinidia eriantha*.

Additional file 13. The results of niche identity test between West group and East group in subtropical China. The histogram indicate the randomized distributions of Warren et al.'s I and Schoener's D and the arrow indicates the observed values of I and D . The x-axis indicates values of I and D , and the y-axis indicates the number of randomizations.

Additional file 14. cpDNA haplotype with its Gene Bank accession number.

Acknowledgements

We are grateful to Ping Tang, Chunyan Liu, Zhengwang Jiang, Zuozhou Li, Puxin Gao, Yueyuan Liu, Weihong Bao and Nan Lin for their assistance in the field observation and collecting samples. We thank the China Scholarship Council (CSC) (No. [2018]3101) for support to Rui Guo.

Authors' contributions

RG and XHY designed the research and collected the samples; RG and YHZ performed the experiment, analysis and wrote the manuscript; all authors revised the manuscript. The author(s) read and approved the final manuscript.

Funding

This work was supported by the Natural Scientific Foundation of China [31770374] and the Strategic Priority Research Program of Chinese Academy of Sciences [XDA20050203].

Availability of data and materials

All haplotype sequences are deposited in GenBank (accession numbers: MN974299-974370; MW387118-387126) (Additional file 14).

Declarations

Ethics approval and consent to participate

XHY undertook the formal identification of the plant material used in our study. The authors have complied with all relevant institutional and national guidelines and legislation in experimental research and field studies on plants, including the collection of plant materials for this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, The Chinese Academy of Sciences, Wuhan 430074, Hubei, China. ²Center of Conservation Biology, Core Botanical Gardens, The Chinese Academy of Sciences, Wuhan 430074, Hubei, China. ³College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. ⁴Department of Plant Sciences, University of Cambridge, Tennis Court Road, Cambridge CB2 3EA, UK. ⁵College of Life and Environmental Sciences, Wenzhou University, Wenzhou 325035, China. ⁶School of Integrative Plant Science, Section of Plant Biology and the L.H. Bailey Hortorium, Cornell University, Ithaca, NY 14853, USA. ⁷BTI Computational Biology Center, Boyce Thompson Institute, Ithaca, NY 14853, USA.

Received: 15 December 2021 Accepted: 10 February 2022

Published online: 28 February 2022

References

- Stewart JR, Lister AM, Barnes I, Dalén L. Refugia revisited: individualistic responses of species in space and time. *Proc R Soc B Biol Sci*. 2010;277:661–71.
- Keppel G, Van Niel KP, Wardell-Johnson GW, Yates C, Byrne M, Mucina L, et al. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Glob Ecol Biogeogr*. 2012;21:393–404.
- Heal B. The Arctic is an ecosystem. 2000. <http://www.thearctic.is>.
- Avise JC. *Phylogeography: the history and formation of species*. Cambridge: Harvard University Press; 2000.
- Zhang JJ, Li ZZ, Pritsch PW, Tian H, Yang AH, Yao XH. Phylogeography and genetic structure of a tertiary relict tree species, *Tapiscia sinensis* (Tapisciaceae): implications for conservation. *Ann Bot*. 2015;116:727–37.
- Kou YX, Cheng SM, Tian S, Li B, Fan DM, Chen YJ, et al. 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;43:351–60.
- Wang YH, Comes HP, Cao YN, Guo R, Mao YR, Qiu YX. Quaternary climate change drives allo-peripatric speciation and refugial divergence in the *Dyosma versipellis-pleiantha* complex from different forest types in China. *Sci Rep*. 2017;7:40261.
- Tian S, Kou YX, Zhang ZR, Yuan L, Li DR, López-Pujol J, et al. Phylogeography of *Eomecon chionantha* in subtropical China: the dual roles of the Nanling Mountains as a glacial refugium and a dispersal corridor. *BMC Evol Biol*. 2018;18:20.
- Ye JW, Li DZ, Hampe A. Differential quaternary dynamics of evergreen broadleaved forests in subtropical China revealed by phylogeography of *Lindera aggregata* (Lauraceae). *J Biogeogr*. 2019;46:1112–23.
- Wang Z, Zhong CH, Li DW, Yan CL, Yao XH, Li ZZ. Cytotype distribution and chloroplast phylogeography of the *Actinidia chinensis* complex. *BMC Plant Biol*. 2021;21:325.
- Wang YH, Jiang WM, Comes HP, Hu FS, Qiu YX, Fu CX. 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;206:852–67.
- Zhang YH, Wang IJ, Comes HP, Peng H, Qiu YX. Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure. *Sci Rep*. 2016;6:24041.
- Fan DM, Sun ZX, Li B, Kou YX, Hodel RGJ, Jin ZN, et al. Dispersal corridors for plant species in the Poyang Lake Basin of southeast China identified by integration of phylogeographic and geospatial data. *Ecol Evol*. 2017;7:5140–8.
- Yu XQ, Gao LM, Soltis DE, Soltis PS, Yang JB, Fang L, et al. Insights into the historical assembly of east Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytol*. 2017;125:2135–48.
- Chen YS, Deng T, Zhou Z, Sun H. Is the east Asian flora ancient or not? *Natl Sci Rev*. 2017;5:920–32.

16. Chen LZ, Sun H, Guo K. Flora and vegetation geography of China. Beijing: Science Press; 2015.
17. Wu ZY, Wang X, Liu F, Zhou F. Vegetation of China. Beijing: Science Press; 1980.
18. Sun X, Chen Y. Palynological records of the last 11 000 yr in China. *Quat Sci Rev*. 1991;10:537–45.
19. Zhou Y, Qiu G, Guo D. Changes of permafrost in China during quaternary. In: Liu TS, editor. Quaternary geology and environment in China. Beijing: Science Press; 1991. p. 86–94.
20. Shi MM, Michalski SG, Welk E, Chen XY, Durka W. Phylogeography of a widespread Asian subtropical tree: genetic east–west differentiation and climate envelope modelling suggest multiple glacial refugia. *J Biogeogr*. 2014;41:1710–20.
21. Fan DM, Hu W, Li B, Morris AB, Zheng M, Soltis DE, et al. Idiosyncratic responses of evergreen broad-leaved forest constituents in China to the late Quaternary climate changes. *Sci Rep*. 2016;6:31044.
22. 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;16:23–38.
23. Tian S, Lei SQ, Hu W, Deng LL, Li B, Meng QL, et al. Repeated range expansions and inter-/postglacial recolonization routes of *Sargentodoxa cuneata* (Oliv.) Rehd. et Wils. (Lardizabalaceae) in subtropical China revealed by chloroplast phylogeography. *Mol Phylogenet Evol*. 2015;85:238–46.
24. Gao YD, Zhang Y, Gao XF, Zhu ZM. Pleistocene glaciations, demographic expansion and subsequent isolation promoted morphological heterogeneity: a phylogeographic study of the alpine *Rosa sericea* complex (Rosaceae). *Sci Rep*. 2015;5:11698.
25. Řičanová Š, Koshev Y, Řičan O, Čosić N, Čirović D, Sedláček F, et al. Multilocus phylogeography of the European ground squirrel: cryptic interglacial refugia of continental climate in Europe. *Mol Ecol*. 2013;22:4256–69.
26. Tang YS, Shi W, Zeng WH, Zheng WY, Cao KF. Floristic composition and phylogenetic diversity of climbing plants in natural forests across Guangxi. *Sheng Tai Xue Bao*. 2018;38:8750–7.
27. Li JQ, Li XW, Soejarto DD. A revision of the genus *Actinidia* from China. *Acta Hort*. 2007;753:41–4.
28. Huang H. Genetic resources. In: Testolin R, Huang HW, Ferguson AR, editors. The kiwifruit genome. Berlin/Heidelberg: Springer International; 2016.
29. Liu YF, Liu YL, Huang H. Genetic variation and natural hybridization among sympatric *Actinidia* species and the implications for introgression breeding of kiwifruit. *Tree Genet Genomes*. 2010;6:801–13.
30. Tang JL, Huang CH, Wu H, Lang BB, Xu XB. Genetic diversity of wild *Actinidia eriantha* germplasm based on fruit traits and SSR markers. *Yuan Yi Xue Bao*. 2014;41:1198–2106.
31. Guo R, Landis JB, Moore MJ, Meng AP, Jian SG, Yao XH, et al. Development and application of transcriptome-derived microsatellites in *Actinidia eriantha* (Actinidiaceae). *Front Plant Sci*. 2017;8:1383.
32. Nybom H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol Ecol*. 2004;13:1143–55.
33. Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, et al. Analysis of SSRs derived from grape ESTs. *Theor Appl Genet*. 2000;100:723–6.
34. Gonneau C, Noret N, Godé C, Frérot H, Sirguyet C, Sterckeman T, et al. Demographic history of the trace metal hyperaccumulator *Noccaea caerulescens* (J. Presl and C. Presl) F. K. Mey. in Western Europe. *Mol Ecol*. 2017;26:904–22.
35. Tsuda Y, Semerikov V, Sebastiani F, Vendramin GG, Lascoux M. Multispecies genetic structure and hybridization in the *Betula* genus across Eurasia. *Mol Ecol*. 2017;26:589–605.
36. Lewis PO, Crawford DJ. Pleistocene refugium endemics exhibit greater allozymic diversity than widespread congeners in the genus *Polygonella* (Polygonaceae). *Am J Bot*. 1995;82:141–9.
37. Hewitt GM. The genetic legacy of the quaternary ice ages. *Nature*. 2000;405:907–13.
38. Byrne M. Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quat Sci Rev*. 2008;27:2576–85.
39. Gómez A, Lunt DH. Refugia within Refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N, editors. Phylogeography of southern European Refugia. Dordrecht: Springer; 2007.
40. Tian SF, Yasunari T. Climatological aspects and mechanism of spring persistent rains over central China. *J Meteorol Soc Jpn*. 1998;76:57–71.
41. Wu GX, Liu YM, Zhang Q, Duan AM. The influence of mechanical and thermal forcing by the Tibetan plateau on Asian climate. *J Hydrometeorol*. 2007;8:770–89.
42. Wan RJ, Wu GX. Mechanism of the spring persistent rains over southeastern China. *Sci China Ser D*. 2007;50:130–44.
43. Jouzel J, Masson-Delmotte V, Cattani O, Dreyfus G, Falourd S, Hoffmann G, et al. Orbital and millennial Antarctic climate variability over the past 800000 years. *Science*. 2007;317:793–6.
44. Cao YN, Comes HP, Sakaguchi S, Chen LY, Qiu YX. Evolution of East Asia's Arcto-Tertiary relict *Euptelea* (Eupteleaceae) shaped by late Neogene vicariance and quaternary climate change. *BMC Evol Biol*. 2016;16:66.
45. Shafer A, Wolf JB. Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecol Lett*. 2013;16:940–50.
46. Riddle BR. Comparative phylogeography clarifies the complexity and problems of continental distribution that drove AR Wallace to favor islands. *PNAS*. 2016;113:7970–7.
47. Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Mol Ecol*. 2013;22:5983–99.
48. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochem Bull*. 1987;19:11–5.
49. Chat J, Chalak L, Petit RJ. Strict paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in intraspecific crosses of kiwifruit. *Theor Appl Genet*. 1999;99:314–22.
50. Weising K, Fung RWM, Keeling DJ, Atkinson RG, Gardner RC. Characterization of microsatellites from *Actinidia chinensis*. *Mol Breed*. 1996;2:117–31.
51. Huang WG, Cipriani G, Morgante M, Testolin R. Microsatellite DNA in *Actinidia chinensis*: isolation, characterisation, and homology in related species. *Theor Appl Genet*. 1998;97:1269–78.
52. Fraser LG, Harvey CF, Crowhurst RN, De Silva HN. EST-derived microsatellites from *Actinidia* species and their potential for mapping. *Theor Appl Genet*. 2004;108:1010–6.
53. Dieringer D, Schlotterer C. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes*. 2003;3:167–9.
54. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–9.
55. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25:1451–2.
56. Dupanloup I, Schneider S, Excoffier L. A simulated annealing approach to define the genetic structure of populations. *Mol Ecol*. 2002;11:2571–81.
57. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992;131:479–91.
58. 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.
59. Clement M, Posada D, Crandall KA. TCS: a computer program to estimate gene genealogies. *Mol Ecol*. 2000;9:1657–9.
60. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, et al. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 2014;10:e1003537.
61. Durraba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 2012;9:772.
62. Liu YF, Li DW, Zhang Q, Song C, Zhong CH, Zhang XD, 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;215:877–90.
63. Rambaut A, Suchard MA, Xie D, Drummond AJ. Tracer v1.6. MCMC trace analyses tool. 2014. Available from <http://beast.bio.ed.ac.uk/Tracer>.
64. Heled J, Drummond AJ. Bayesian inference of population size history from multiple loci. *BMC Evol Biol*. 2008;8:289.
65. Harpending HC. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum Biol*. 1994;66:591–600.
66. Rogers AR. Genetic evidence for a pleistocene population explosion. *Evolution*. 1995;49:608–15.
67. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123:585–95.

68. Fu YX. Statistical tests of neutrality of mutations against population growth, hitch-hiking, and background selection. *Genetics*. 1997;147:915–25.
69. Yu Y, Harris AJ, Blair C, He XJ. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Mol Phylogenet Evol*. 2015;87:46–9.
70. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol*. 1981;17:368–76.
71. Stucki S, Orozco-terWengel P, Forester BR, Duruz S, Colli L, Masembe C, et al. High performance computation of landscape genomic models integrating local indices of spatial association. *Mol Ecol Resour*. 2017;17:1072–89.
72. Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, et al. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am Nat*. 2016;188:379–97.
73. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000;25:25–9.
74. Kalinowski ST, Taper ML, Marshall TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*. 2007;16:1099–106.
75. Goudet J. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). 2001. Available online at: <http://www2.unil.ch/popgen/softwares/fstat.htm>.
76. Chapuis MP, Estoup A. Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol*. 2007;24:621–31.
77. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*. 1978;89:583–90.
78. Peakall R, Smouse PE. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol Ecol Notes*. 2006;6:288–95.
79. Vandergast AG, Perry WM, Lugo RV, Hathaway S. Genetic landscapes GIS toolbox: tools to map patterns of genetic divergence and diversity. *Mol Ecol Resour*. 2011;11:158–61.
80. Nei M, Tajima F, Tateno Y. Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol*. 1983;19:927–43.
81. Takezaki N, Nei M, Tamura K. Poptree 2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Mol Biol Evol*. 2010;27:747–52.
82. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
83. 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.
84. 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.
85. Beerli P, Palczewski M. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*. 2010;185:313–26.
86. Schlötterer C. Evolutionary dynamics of microsatellite DNA. *Chromosoma*. 2000;109:365–71.
87. Wilson GA, Rannala B. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*. 2003;163:1177–91.
88. Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, et al. 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.
89. Cornuet JM, Ravigné V, Estoup A. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*. 2010;11:401.
90. 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.
91. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al. Vegan: community ecology package. R package version 2.3–4. 2016.
92. Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*. 1997;145:1219–28.
93. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. *Ecol Model*. 2006;190:231–59.
94. Warren DL, Glor RE, Turelli M. ENM tools: a toolbox for comparative studies of environmental niche models. *Ecography*. 2010;33:607–11.
95. Warren DL, Glor RE, Turelli M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*. 2008;62:2868–83.
96. Schoener TW. The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology*. 1968;49:704–26.
97. Brown JL. SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol Evol*. 2014;5:694–700.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

