

SCIENTIFIC REPORTS



OPEN

Quaternary climate change drives allo-peripatric speciation and refugial divergence in the *Dysosma versipellis-pleiantha* complex from different forest types in China

Received: 20 September 2016

Accepted: 05 December 2016

Published: 11 January 2017

Yi-Han Wang¹, Hans Peter Comes², Ya-Nan Cao¹, Rui Guo¹, Yun-Rui Mao¹ & Ying-Xiong Qiu¹

Subtropical China harbours the world's most diverse temperate flora, but little is known about the roles of geographical and eco-climatic factors underlying the region's exceptionally high levels of species diversity and endemism. Here we address this key question by investigating the spatio-temporal and ecological processes of divergence within the *Dysosma versipellis-pleiantha* species complex, endemic to subtropical China. Our cpDNA phylogeny showed that this monophyletic group of understory herbs is derived from a Late Pliocene ancestor of the Qinghai-Tibetan Plateau (QTP)/Southwest China. Genetic and ENM data in conjunction with niche differentiation analyses support that the early divergence of *D. versipellis* and *D. pleiantha* proceeded through allo-peripatric speciation, possibly triggered by Early Pleistocene climate change, while subsequent climate-induced cycles of range contractions/expansions enhanced the eco-geographical isolation of both taxa. Furthermore, modelling of population-genetic data indicated that major lineage divergences within *D. versipellis* likely resulted from long-term allopatric population isolation in multiple localized refugia over the last glacial/interglacial periods, and which in turn fostered endemic species formation (*D. difformis*, *D. majoensis*) from within *D. versipellis* in Southwest China. These findings point to an overriding role of Quaternary climate change in triggering essentially allopatric (incipient) speciation in this group of forest-restricted plant species in subtropical China.

Subtropical (Central/South/East) China (c. 22°N–30/33°N), the focal region of the present study, harbours the world's most diverse temperate flora¹, as mainly found in lower-elevation warm-temperate evergreen (WTE) vs. montane warm-temperate deciduous (WTD) forest (c. 300–800 m vs. c. 800–2000 m)^{2,3}. However, little is known about the times and processes of diversification underlying the region's exceptionally high species diversity and endemism. Recent phylogeographic studies of various plant species (e.g. trees, shrubs, herbs) representing either forest biome have provided evidence for long-term allopatric population isolation (viz. vicariance) across multiple localized refugia (e.g. WTD^{4–9}; WTE^{10–15}). For some of them (e.g. *Emmenopterys henryi*⁸), intraspecific genetic divergence has also been demonstrated to partly result from local adaptation to different climatic environments. However, as these recent efforts have all focussed on single species, it remains unclear to what extent geographical and/or eco-climatic factors have driven speciation among closely related plant taxa that vary in their geographic distributions and/or forest biome affinities across subtropical China.

Here we examine the evolutionary history of four closely related species of the '*Dysosma versipellis-pleiantha*' complex (*sensu*¹⁶). Our preliminary analyses of nuclear ribosomal (nr) and chloroplast (cp) DNA¹⁷ placed this species complex as sister group to the rest of the genus, comprised of three high-elevation species of cool-temperate/subalpine forest (c. 1700–3900 m) from Yunnan (Southwest China) and the south-eastern slopes of the Qinghai-Tibetan Plateau (QTP); however, these analyses failed to further resolve relationships within the species complex. Notably, its four recognized species differ in their degree of range overlap and/or type of forest habitat

¹Key Laboratory of Conservation Biology for Endangered Wildlife of the Ministry of Education, and Laboratory of Systematic & Evolutionary Botany and Biodiversity, College of Life Sciences, Zhejiang University, Hangzhou 310058, China. ²Department of Ecology and Evolution, Salzburg University, A-5020 Salzburg, Austria. Correspondence and requests for materials should be addressed to Y.-X.Q. (email: qyxhero@zju.edu.cn)

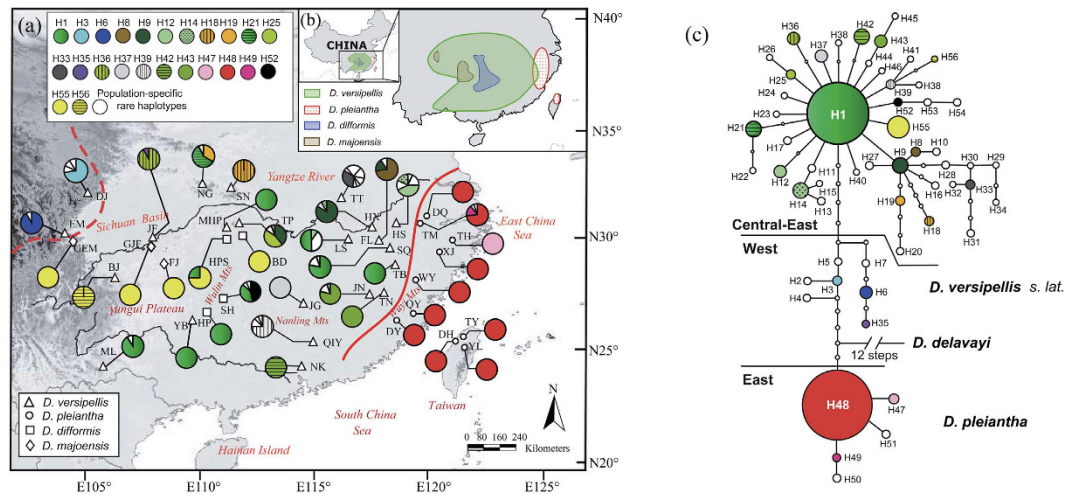


Figure 1. (a) Geographic distribution of the 56 chloroplast (cp) DNA (*trnL-trnF*, *trnL-ndhJ*, *trnS-trnFM*) haplotypes (H1–H56) detected in the *Dysosma versipellis-pleiantha* complex, created in ILLUSTRATOR v15.0 (<http://www.adobe.com/products/illustrator.html>) (see Supplementary Table S2 for population codes). Red dashed and solid lines denote the major western vs. central-east split within *D. versipellis s. lat.* (i.e. *D. versipellis*, *D. difformis*, *D. majoensis*) and the divergence between the latter and *D. pleiantha* further east, respectively, as identified by phylogenetic tree and network analyses. (b) Geographic ranges of *D. versipellis*, *D. pleiantha*, *D. difformis*, and *D. majoensis*¹⁸, created in ILLUSTRATOR. The base map was drawn using ARCGIS v.9.3 (ESRI, Redlands, CA, USA). (c) Ninety-five percent statistical parsimony network of the 56 cpDNA haplotypes identified in the species complex generated in TSC v1.21 (<http://darwin.uvigo.es/software/tcs.html>). The small open circles represent missing haplotypes. The size of circles corresponds to the haplotype frequency.

across subtropical China (Fig. 1a,b)^{16,18}. Two of them, *D. versipellis* (Hance) M. Cheng ex Ying and *D. pleiantha* (Hance) Woodson, form a parapatric species pair with abutting ranges in East China; the former occurs in isolated stands of montane WTD forest [c. 800–2000 (2700 m)] throughout Central-East China, whereas *D. pleiantha* is associated with WTE forest habitats in the hilly/coastal areas of East China (300–800 m) and the mountains of Taiwan (1000–2500 m). By contrast, *D. difformis* (Hemsley & E.H. Wilson) T.H. Wang ex T.S. Ying and *D. majoensis* (Gagnepain) M. Hiroe are rare endemics with proximate but non-overlapping ranges in the WTD forests of Central-Southwest China, where both are broadly sympatric (but usually not locally admixed) with *D. versipellis* (Fig. 1b). Despite clear taxonomic arguments for the recognition of four species within the complex^{16,18}, the possibility of hybridization has long been suspected from morphological and/or distributional considerations, e.g. especially between *D. versipellis* and both *D. pleiantha* and the subalpine *D. delavayi* (Franchet) Hu¹⁶.

In this study, we combined phylogeographic analyses using both nuclear and cpDNA data with cpDNA-based phylogenetic and dating approaches at the genus level, as well as Approximate Bayesian Computation (ABC), (palaeo-)climatic data and niche identity tests, to elucidate the spatial-temporal and ecological processes of divergence within the species complex that eco-geographically varies across the two major forest biomes of subtropical China. The specific aims of this study were: (1) to evaluate the phylogenetic relationships and divergence times of the four focal species and their intraspecific lineages; (2) to compare patterns of genetic diversity and structure among these four species; and (3) to infer the phylogeographic and demographic history of the species complex, including levels of admixture and/or hybridization.

Results

Chloroplast haplotype diversity and population structure. The concatenated cpDNA sequences (*trnL-trnF*, *trnL-ndhJ*, *trnS-trnFM*), surveyed across 452 individuals of the *D. versipellis-pleiantha* complex plus 10 accessions representing the five outgroup species, were aligned with a total length of 2604 bp, including 37 substitutions and 20 indels (1–23 bp). Within the species complex, these polymorphisms identified 56 haplotypes (Supplementary Table S1). Of those, 46 were specific to *D. versipellis* (H2–H46, H56), five to *D. pleiantha* (H47–H51) and three to *D. difformis* (H52–H54), while *D. majoensis* was fixed for H55 (Fig. 1a). H55 was shared between *D. difformis* and *D. majoensis*, and H1, the most widespread and common haplotype of *D. versipellis*, was also found in *D. difformis*. For the species complex as a whole, the data revealed high levels of haplotype diversity ($h_T = 0.929$) and nucleotide diversity ($\pi_T = 3.58 \times 10^{-3}$). The widespread *D. versipellis* had on average only slightly higher levels of within-population diversity ($h_S = 0.300$; $\pi_S = 0.24 \times 10^{-3}$) than the geographically more restricted *D. difformis* ($h_S = 0.289$; $\pi_S = 0.14 \times 10^{-3}$), whereas *D. pleiantha* was almost fixed for H48 (Supplementary Table S2; Fig. 1a). Hierarchical AMOVA apportioned 62% of the total cpDNA variance among the four species, whereas separate, non-hierarchical analyses revealed stronger population genetic structure in both *D. versipellis* ($\Phi_{ST} = 0.80$) and *D. difformis* ($\Phi_{ST} = 0.63$) compared to *D. pleiantha* ($\Phi_{ST} = 0.52$) (Table 1). A significant cpDNA phylogeographic structure was indicated for the whole species complex ($N_{ST} = 0.866 > G_{ST} = 0.790$), *D. versipellis* (0.818 vs. 0.611) and *D. difformis* (0.865 vs. 0.728) (all $P < 0.05$), but not for *D. pleiantha* (0.204 vs. 0.161, $P > 0.05$).

Source of variation	cpDNA			EST-SSR (All 15 loci/nine neutral loci)		
	d.f.	Percentage of variation (%)	Φ -statistics	d.f.	Percentage of variation (%)	R-statistics
<i>D. versipellis-pleiantha</i> complex						
Among species	3	62.45	$\Phi_{CT} = 0.62$	3/3	36.53/14.09	$R_{CT} = 0.365/0.141$
Among populations within species	34	29.97	$\Phi_{SC} = 0.80$	36/36	34.54/46.65	$R_{SC} = 0.544/0.543$
Within populations	425	7.58	$\Phi_{ST} = 0.92$	1114/1114	28.93/39.27	$R_{ST} = 0.711/0.607$
<i>D. versipellis-pleiantha</i> complex						
Among cpDNA lineages	2	74.91	$\Phi_{CT} = 0.75$			
Among populations within lineages	34	18.67	$\Phi_{SC} = 0.74$			
Within populations	425	6.42	$\Phi_{ST} = 0.94$			
<i>Dysosma versipellis</i>						
Among populations	20	80.40	$\Phi_{ST} = 0.80$	21/21	61.60/61.69	$R_{ST} = 0.616/0.617$
Within populations	269	19.60		679/679	38.40/38.30	
<i>Dysosma pleiantha</i>						
Among populations	9	51.77	$\Phi_{ST} = 0.52$	10/10	28.97/24.16	$R_{ST} = 0.290/0.242$
Within populations	114	48.23		345/345	71.03/75.84	
<i>Dysosma difformis</i>						
Among populations	3	62.91	$\Phi_{ST} = 0.63$	3/3	37.64/29.45	$R_{ST} = 0.376/0.294$
Within populations	31	37.09		65/65	62.35/70.55	
<i>Dysosma majoensis</i>						
Among populations				2/2	71.90/72.55	$R_{ST} = 0.719/0.725$
Within populations				25/25	28.10/27.45	

Table 1. Results of analyses of molecular variance (AMOVAs) based on cpDNA haplotype data and EST microsatellite allele frequencies for species and populations of the *D. versipellis-pleiantha* complex. All levels of variation were significant. Note that *D. majoensis* was proved invariable at the cpDNA level.

Phylogenetic haplotype relationships and molecular dating. The cpDNA tree topologies obtained from Bayesian inference (BI) (Fig. 2) and maximum likelihood (ML, not shown) supported the monophyly of both *Dysosma* (PP = 1, ML = 95%) and the *D. versipellis-pleiantha* complex (PP = 1, ML = 93%), and resolved *D. delavayi* as most likely sister to the latter group (PP = 0.95, ML = 83%), while basal relationships between *D. aurantiocaulis* and *D. tsayuensis* remained unresolved. Within the species complex, three geographically structured lineages were resolved: (i) a well-supported, predominant ‘western’ lineage (PP = 0.98, ML = 71%), comprising all haplotypes of *D. versipellis* from the eastern slopes of the QTP (pops. EM, DJ: H2–H7), plus a haplotype (H35) unique to population JF from the Sichuan Basin (Fig. 1a); (ii) a weakly supported ‘central-eastern’ lineage (PP = 0.92, ML = 42%), containing all the remaining haplotypes of *D. versipellis*, and those of *D. difformis* and *D. majoensis* from Central-Southwest China; and (iii) a strongly-supported ‘eastern’ lineage (PP = 1, ML = 91%), consisting of all *D. pleiantha* haplotypes (Fig. 2). Although the ‘western’ and ‘eastern’ lineages appeared to be sisters, this relationship received very low support (PP = 0.94, ML = 56%). Individual mean age estimates for relevant nodes and their 95% highest posterior densities (HPDs) are provided for the different coalescent-type models in Supplementary Table S3. Accordingly, the *D. versipellis-pleiantha* complex likely originated in the Late Pliocene at *c.* 3.33 million yr ago (Ma) (constant size model) vs. *c.* 2.66 Ma (expansion growth model) (node ‘B’), and started to diversify in the Early Pleistocene (*c.* 2.16 vs 1.66 Ma) (node ‘C’), whereas the crown ages of the main lineages fell into the early-to-mid Pleistocene (central-east: *c.* 1.36 vs. 1.02 Ma, node ‘E’; west: *c.* 0.91 vs. 0.70 Ma, node ‘F’; east/*D. pleiantha*: *c.* 0.59 vs. 0.52 Ma, node ‘G’) (Supplementary Table S3).

In the rooted TCS network (Fig. 1c), all haplotypes of *D. pleiantha* were recovered as sister clade to the remainder, separated by five mutational steps. Within this latter clade (hereafter named ‘*D. versipellis s. lat.*’), haplotypes of the ‘western’ vs. ‘central-eastern’ lineages formed interior (ancestral) vs. tip (derived) subclades, again, separated by five steps. Notably, the low frequency haplotypes of both *D. difformis* (H52–H54) and *D. majoensis* (H55) appeared to be mutational derivatives of the most common and widespread haplotype of *D. versipellis* (H1) as part of the ‘central-east’ subclade.

Demographic history. The mismatch distribution for cpDNA haplotypes of the ‘western’ lineage of *D. versipellis s. lat.* was bimodal, and thus differed strongly from that predicted under the spatial and demographic expansion models (Supplementary Fig. S1). This difference was also registered by non-significant values of both Fu’s F_S and Tajima’s D (Table 2). By contrast, both parameter values were significantly negative for the ‘central-east’ lineage of *D. versipellis s. lat.*, indicative of population expansion, and the same applied to *D. pleiantha* in terms of F_S (Table 2). Moreover, for either lineage, the MDAs gave unimodal graphs as well as non-significant SSD and H_{Reg} values under both the spatial and demographic models (Supplementary Fig. S1; Table 2). Based on the corresponding τ values and the assumed substitution rate of 1.3×10^{-9} s/s/y, we calculated a mid-Pleistocene expansion time for the ‘central-east’ lineage, regardless of the model used [demographic: *c.* 0.48 Ma (95% CI: 0.21–0.69 Ma); spatial: *c.* 0.36 (0.20–0.61) Ma]. For *D. pleiantha*, we obtained a similar timing under the

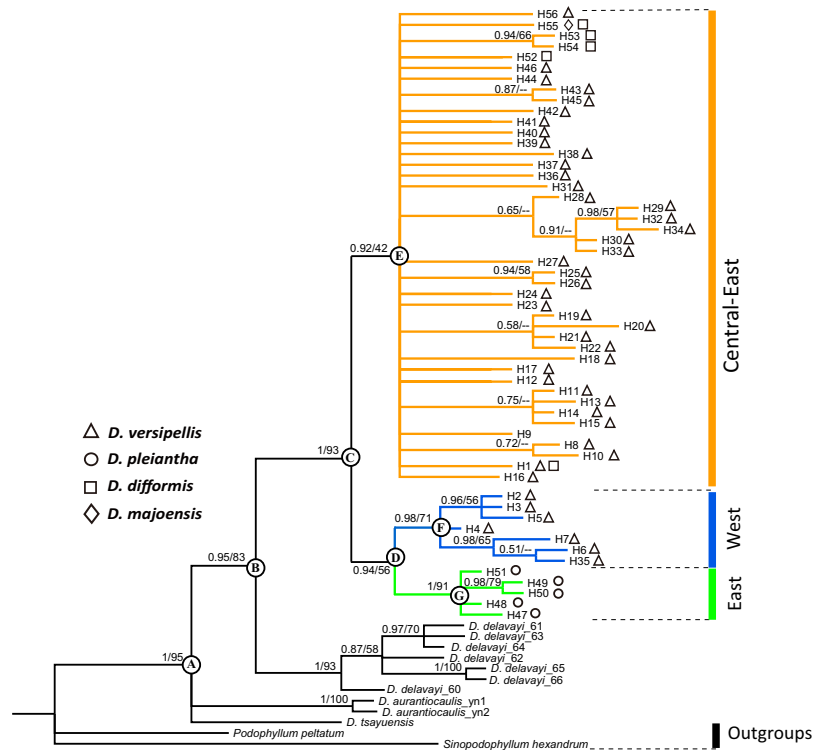


Figure 2. Bayesian inference (BI) phylogeny of the *Dysosma versipellis-pleiantha* complex based on cpDNA (*trnL-trnF*, *trnL-ndh*), *trnS-trnM* sequences with the remaining *Dysosma* species treated as part of the ingroup and *Podophyllum peltatum* and *Sinopodophyllum hexandrum* as outgroups. Posterior probabilities (PP > 0.50) and maximum likelihood (ML) bootstrap values (>50%) are sequentially indicated above the branches. Colored branches identify major haplotype lineages (west, central-east, east) within the species complex. Symbols following the haplotype numbers indicate the species bearing this haplotype. Nodes of interest are marked as A–G, while the corresponding BEAST-derived age estimates (including their 95% HPD intervals) are shown in Supplementary Table S3.

Model	Lineage	Parameter (τ)	Expansion time (t) in yr BP	SSD	P	$HRag$	P	Fu's F_s	P	Tajima's D	P
Demographic expansion	West	4.910 (0.207–8.426)	NC	0.1349	0.05	0.368	0.01	-0.19	0.47	0.5	0.71
	Central-east	3.350 (1.455–4.800)	482,750 (209,672–691,702)	0.0005	0.92	0.015	0.87	-22.3	0.00	-1.59	0.02
	East	3.000 (0.512–3.000)	432,314 (73,782–432,314)	0.0000	0.47	0.562	0.67	-3.06	0.02	-1.24	0.07
Spatial expansion	West	4.270 (1.791–7.839)	NC	0.0985	0.08	0.368	0.25	-0.19	0.45	0.50	0.74
	Central-east	2.474 (1.402–4.207)	356,515 (202,034–606,248)	0.0004	0.96	0.015	0.90	-22.3	0.00	-1.59	0.02
	East	0.255 (0.001–1.035)	36,747 (114–149,148)	0.0000	0.55	0.562	0.66	-3.06	0.02	-1.24	0.09

Table 2. Mismatch distribution analysis (MDA) of the three cpDNA lineages of the *D. versipellis-pleiantha* complex (west, central-east: *D. versipellis* s. lat.; east: *D. pleiantha*) for pure demographic and spatial expansion models, tested with the sum of squared deviations (SSD) and Harpending's⁷⁹ raggedness index ($HRag$) in ARLEQUIN. See Supplementary Fig. S1 for illustrations of mismatch distributions. Also shown are results of selective neutrality tests (Fu's F_s , Tajima's D). For the F_s test, $P = 0.02$ is considered to be significant at the $\alpha = 0.05$ level^{59,80}. NC, not calculated.

demographic model [c. 0.43 (0.07–0.43) Ma], whereas the spatial model was more indicative of expansion during the Late Pleistocene/Holocene [c. 0.04 (0.0001–0.15) Ma].

EST-SSR diversity, population structure and migration/gene flow. A total of 138 alleles were identified over the 15 EST-SSR loci among the 577 individuals (40 populations) surveyed. For each locus, the frequency of null alleles detected by FREENA was lower than the threshold ($\nu = 0.15$) across all analyzed populations. There was no evidence for a systematic departure from HWE at particular loci, or a significant LD detectable among loci (data not shown). Using different outlier detection methods (see Supplementary Table S4), congruent evidence was found at each of three loci for either positive (EDV-46/52/119) or balancing selection (EDV-60/67/82).

At the species level, measures of genetic diversity derived from all 15 loci were highest in *D. versipellis* (e.g. $H_E = 0.618$), followed by *D. difformis* ($H_E = 0.610$), *D. pleiantha* ($H_E = 0.473$), and *D. majoensis* ($H_E = 0.447$;

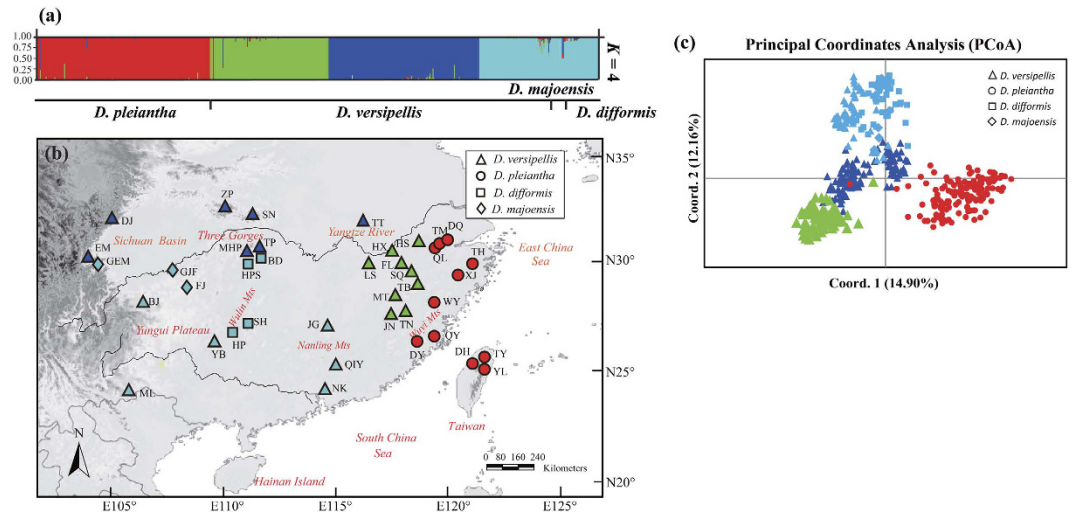


Figure 3. (a) Histogram of the assignment test for 40 populations (577 individuals) of the *Dysosma versipellis-pleiantha* complex based on genetic variation at 15 EST-SSR loci using STRUCTURE v2.3.4 (<http://pritchardlab.stanford.edu/structure.html>). Each vertical bar represents one individual and its probability of membership for each of the $K = 4$ clusters. (b) Geographic distribution of the four STRUCTURE clusters within and among populations of the species complex created in ILLUSTRATOR v15.0 (<http://www.adobe.com/products/illustrator.html>). The symbol next to each sampling locality identifies the respective *Dysosma* species, while the filled color represents the cluster assigned to that population (population codes are identified in Supplementary Table S2). The base map was drawn using ARCGIS v.9.3 (ESRI, Redlands, CA, USA). (c) Principal coordinates analysis (PCoA) of the 577 individuals from the four species of the complex based on EST-SSR variation (15 loci) using ARLEQUIN v3.5 (<http://cmpg.unibe.ch/software/arlequin35/>).

Supplementary Table S1). In the hierarchical AMOVA, 36.53% of the total EST-SSR variation was distributed among the four species ($R_{CT} = 0.365$), 34.54% was explained by variation among populations within species ($R_{SC} = 0.544$), and 28.93% was apportioned within populations ($R_{ST} = 0.711$; Table 1). Non-hierarchical analyses revealed stronger population genetic structure in *D. majoriensis* ($R_{ST} = 0.72$) and *D. versipellis* ($R_{ST} = 0.62$), compared to *D. difformis* ($R_{ST} = 0.38$) and *D. pleiantha* ($R_{ST} = 0.29$) (Table 1).

In the STRUCTURE analysis of the entire species complex using all 15 loci, $\ln P(D)$ progressively increased with K , whereas the ad hoc statistic ΔK showed the highest likelihood at $K = 4$ (Supplementary Fig. S2). Based on the latter model, all samples of *D. pleiantha* formed a distinct cluster (red in Fig. 3a,b), while those of *D. versipellis* segregated into three regional clusters, with populations originating from (i) both the west and the north (dark blue); (ii) the south (light blue); and (iii) the east (green). All samples of *D. difformis* and *D. majoriensis* were assigned to the southern cluster of *D. versipellis* (Fig. 3a,b). Similarly, in the corresponding PCoA, the first axis distinguished individuals of *D. pleiantha* from all others, whereas individuals of *D. difformis* and *D. majoriensis* broadly overlapped with the southern cluster of *D. versipellis* in genetic space (Fig. 3c). After removing the six outlier loci, ΔK favoured a model with $K = 3$. Running STRUCTURE and using $K = 3$, *D. pleiantha* and eastern populations of *D. versipellis* were assigned to two distinct gene pools, whereas the remainder made up one single gene pool (see Supplementary Fig. S3). However, with $K = 4$, results were consistent with the clustering patterns inferred using all 15 loci (Supplementary Fig. S3).

The MIGRATE analysis using all 15 loci revealed all pairwise estimates of migration/gene flow (M) among *D. pleiantha*, *D. difformis*, *D. majoriensis* and the three regional clusters of *D. versipellis* to be low, ranging from 0.351 (*D. difformis* to west-northern cluster of *D. versipellis*) to 2.160 (eastern to west-northern cluster of *D. versipellis*) (Table 3). Basically the same results were obtained after removing the outlier loci (Table 3). Overall, these STRUCTURE, PCoA and MIGRATE analyses indicate only relatively minor effects of outlier loci on the overall population structure and gene flow of this species complex.

ABC-based inferences of species/cluster divergence. In the DIYABC analysis of the entire species complex based on the nine neutral EST-SSR loci, the highest posterior probabilities derived from the direct estimate approach were found for scenarios 1 (0.31, 95% CI: 0.00–0.71) and 4 (0.29, 0.00–0.69), whereas the values derived from the logistic regression were higher for the former scenario with no overlap of confidence intervals [0.46 (0.45–0.47) vs. 0.27 (0.26–0.28)] (Supplementary Fig. S4). In consequence, we favoured scenario 1, according to which the three regional STRUCTURE clusters of *D. versipellis s. lat.* (west-north, east, south) diverged simultaneously from a common ancestor, rather than any alternative hypothesis of ‘ordered’ relationships among these groups (Supplementary Fig. S5a, Table 4). The DIYABC analysis further estimated that this simultaneous diversification of *D. versipellis s. lat.* occurred during the mid-Pleistocene ($c. 0.59$ Ma, 95% CI: 0.15–1.29 Ma; t_2 in Supplementary Fig. S5a), while its divergence from *D. pleiantha* was dated to the Early Pleistocene [$c. 0.92$ (0.25–2.48) Ma; t_1]. This latter timing broadly concurs with the BEAST-derived crown ages of the entire species complex estimated from cpDNA (2.25 and 1.81 Ma; see above).

	<i>D. pleiantha</i>	<i>D. versipellis</i> East (E)	<i>D. versipellis</i> West-north (WN)	<i>D. versipellis</i> South (S)	<i>D. majoensis</i>	<i>D. difformis</i>
<i>D. pleiantha</i>		1.343 (1.165–1.540)	1.887 (1.454–2.125)	1.018 (0.862–1.348)	0.457 (0.356–0.575)	0.377 (0.286–0.485)
		2.617 (2.232–2.958)	1.337 (1.078–1.580)	0.844 (0.658–1.044)	0.851 (0.656–1.047)	0.071 (0.004–0.262)
<i>D. versipellis</i> E	1.887 (1.659–2.136)		0.808 (0.586–0.974)	0.626 (0.499–0.773)	0.483 (0.341–0.615)	0.351 (0.252–0.465)
	1.479 (1.211–1.759)		1.196 (0.984–1.471)	0.398 (0.279–0.560)	0.625 (0.475–0.804)	0.304 (0.134–0.434)
<i>D. versipellis</i> WN	1.531 (1.342–1.737)	2.160 (1.936–2.401)		1.044 (0.891–1.213)	0.950 (0.804–1.114)	0.443 (0.346–0.601)
	2.495 (2.127–2.827)	1.828 (1.568–2.162)		1.086 (0.836–1.310)	0.578 (0.411–0.745)	0.411 (0.295–0.559)
<i>D. versipellis</i> S	0.572 (0.452–0.710)	1.118 (0.947–1.308)	1.126 (0.863–1.319)		1.069 (0.903–1.255)	1.027 (0.864–1.210)
	0.470 (0.334–0.638)	1.050 (0.830–1.445)	0.589 (0.434–0.790)		1.581 (1.294–1.877)	0.873 (0.683–1.097)
<i>D. majoensis</i>	0.174 (0.098–0.286)	0.526 (0.388–0.692)	0.625 (0.473–0.808)	1.160 (0.944–1.403)		1.042 (0.835–1.279)
	0.144 (0.076–0.244)	0.772 (0.594–0.983)	0.773 (0.508–1.025)	0.246 (0.151–0.391)		0.304 (0.197–0.747)
<i>D. difformis</i>	0.681 (0.525–0.866)	0.550 (0.412–0.717)	1.096 (0.897–1.323)	1.347 (1.123–1.603)	1.483 (1.247–1.748)	
	0.662 (0.474–0.947)	0.669 (0.464–0.898)	0.908 (0.408–1.169)	1.376 (1.100–1.695)	0.875 (0.620–1.133)	

Table 3. Estimates of historical migration rate (M) and 95% confidence intervals (CI) (in parentheses) among species of the *D. versipellis-pleiantha* complex and regional clusters of *D. versipellis* using all 15 (in bold) and mere nine neutral EST-SSR loci. Directionality of gene flow is read from top species/clusters being the source populations, whereas units on the left are the recipient species/clusters.

In the analysis of the southern cluster (Supplementary Fig. S5b; Table 4), scenario 1 provided the best fit to the EST-SSR data [i.e. simultaneous divergence of *D. versipellis* (south), *D. difformis*, *D. majoensis*], with posterior probability values being much higher than for the other three scenarios, regardless of the method used (direct estimate: 0.38, 0.18–0.47; logistic regression: 0.53, 0.52–0.54) (Supplementary Fig. S4). The time of diversification of this southern cluster was dated to the mid-Pleistocene [c. 0.43 (0.07–1.47) Ma; t_4 in Supplementary Fig. S5b].

Ecological niche modelling and niche identity tests. The MAXENT models for *D. versipellis s. lat.* (including *D. difformis*/*D. majoensis*) and *D. pleiantha* had high predictive power and did not overfit the presence data (AUC values = 0.906 ± 0.118 and 0.972 ± 0.041 , respectively). The current potential ranges of either group (defined as modelled suitability ≥ 0.75 ; Fig. 4a,b) were largely consistent with their actual distributions, except for predicted but unsupported occurrences of *D. versipellis s. lat.* in the far east of China, Taiwan, and South Korea (Fig. 4a). During the LGM (c. 21,000 yr BP), *D. versipellis s. lat.* experienced a drastic restriction of its current range in Central-East China (with only scattered areas persisting along the middle-Yangtze River), whereas suitable habitat apparently increased around the Sichuan Basin and areas further south (e.g. Southwest China, Yungui Plateau) (Fig. 4c). By contrast, for *D. pleiantha*, the LGM modelling predicted a more contiguous than currently scattered distribution in East China, extending southward into Taiwan (then connected with Mainland China), plus a higher suitability of areas both west- and eastward, i.e. south of the lower/middle Yangtze and on the then exposed East China Sea basin, respectively (Fig. 4d).

The results of the niche identity tests supported the existence of niche differentiation between the WTD forest-dweller *D. versipellis s. lat.* and its WTE forest counterpart *D. pleiantha*. ENMTOOLS showed that the observed similarity values for I and D (0.818 and 0.582, respectively) were both significantly lower ($P < 0.01$) than the corresponding values expected from the pseudoreplicated datasets in the paired analyses (Supplementary Fig. S6).

Discussion

Our cpDNA phylogeny (Fig. 2) confirms major relationships within *Dysosma* as hypothesized previously on the basis of cpDNA/nrDNA data¹⁷. In both analyses, the *D. versipellis-pleiantha* complex forms a strongly supported monophyletic group, which, in the present analysis, emerges from a paraphyletic and likewise supported grade of the remainder of the genus, comprised of *D. delavayi* (as sister to the complex), *D. aurantiocaulis*, and *D. tsayuensis*. Even though the species complex and the three latter, high-elevation species are not reciprocally monophyletic, they exhibit different morphologies^{16,18} and eco-geographies: while the former group is native to the lowland WTE or montane WTD forests of subtropical China, the latter is endemic to the cool-temperate/subalpine forests of Yunnan and the south-eastern QTP region. Given that the closest relative of *Dysosma*, *Sinopodophyllum hexandrum*, is native to the Himalayan regions^{16,17,19}, we therefore assume that the most recent common ancestor (MRCA) of the *D. versipellis-pleiantha* complex originated in the high-altitude habitats of the QTP region/Southwest China, and then expanded its geographic range across subtropical China. In general, such a niche shift from high-to-lower elevation habitats is thought to be rare²⁰, but is also known from other taxa in our study area (e.g. *Myricaria*²¹) or elsewhere (e.g. East African *Dendrosenecio*²²; neotropical *Hypericum*²³). The divergence of the *D. versipellis-pleiantha* complex from its closest high-altitude relative (*D. delavayi*) most likely occurred during the Late Pliocene (c. 3.33/2.66 Ma; Fig. 2), hence postdating a shift in East Asian climate toward drier conditions and a decrease in summer monsoon rainfall (c. 4–3 Ma^{19,24,25}). These processes, in turn, may have reduced forest habitats²⁶, which could have triggered the isolation of the lowland forms from highland forms.

Within the *D. versipellis-pleiantha* complex, the cpDNA phylogeny (Fig. 2) identified three geographic clades, albeit with different support: a well-supported western clade of *D. versipellis* (mainly distributed east of the QTP), a weakly supported 'central-eastern' clade (distributed in Southwest-Central-East China, and comprised of

Analysis/scenario no.	Parameter	Median	95% lower bound	95% upper bound
(a) Scenario 1				
	NA	137 000	11 300	727 000
	NDP	473 000	187 000	825 000
	NDV _(WN)	593 000	257 000	899 000
	NDV _(E)	316 000	106 000	714 000
	NDV _(S1)	456 000	170 000	838 000
	t_1 (in million years ago)	0.92	0.25	2.48
	t_2 (in million years ago)	0.59	0.15	1.29
	μ	1.99×10^{-6}	1.16×10^{-6}	4.66×10^{-6}
	P	0.38	0.14	0.80
(b) Scenario 1				
	NA	290 000	2 7500	860 000
	NDV _(S2)	459 000	135 000	871 000
	NDD	471 000	139 000	880 000
	NDM	277 000	68 800	755 000
	t_4 (in million years ago)	0.43	0.07	1.47
	μ	2.03×10^{-6}	1.14×10^{-6}	6.45×10^{-6}
	P	0.29	0.12	0.74

Table 4. Descriptions of prior settings and median estimations of posterior distributions of parameters revealed by the DIYABC modeling of the best-fitting scenarios for the diversification history of (a) the three regional EST-SSR (STRUCTURE) clusters of *D. versipellis s. lat.* (west-north, east, south); and (b) the three taxonomic units of its ‘southern cluster’, i.e. *D. versipellis* (south), *D. difformis*, and *D. majoensis*. See Materials and Methods, Supplementary Fig. S5, and Table S5 for details. In (a), NDP, NDV_(WN), NDV_(E), and NDV_(S1) denote the current population sizes of, respectively, *D. pleiantha*, *D. versipellis* west-north cluster, *D. versipellis* east cluster and *D. versipellis* south cluster (including *D. difformis* and *D. majoensis*). In (b), NDV_(S2), NDD, and NDM represent the current population sizes of, respectively, *D. versipellis* south cluster, *D. difformis* and *D. majoensis*. NA represents the population size of the ancestral lineage at the basal node. μ denotes the mutation rate (per generation per locus) for the EST-SSRs. P represents the proportion of multiple step mutations in the generalized stepwise model, GSM.

D. versipellis + *D. difformis* + *D. majoensis*), and a strongly supported eastern clade corresponding to *D. pleiantha*. By contrast, what emerges from our rooted cpDNA haplotype network (Fig. 1c) is striking evidence that *D. pleiantha* is sister to the remainder of the species complex (*D. versipellis s. lat.*). This inference of eco-geographically driven, and possibly adaptive divergence is further strengthened by our niche modeling, which suggests that these taxa occupy significantly different climatic environments (Supplementary Fig. S6). Moreover, in the cpDNA network, the central-eastern clade of *D. versipellis s. lat.* is placed in a derived (tip) position relative to the western (interior) one, and thus most distantly related to *D. pleiantha* (Fig. 1c). Hence, the present-day abutting ranges of *D. versipellis* and *D. pleiantha* in East China are secondary, rather than reflecting primary divergence between adjacent populations of a previously homogeneous species, i.e. a parapatric mode of speciation. We therefore conclude that the main mechanism, which caused the Early Pleistocene [0.92 (0.25–2.48) Ma] separation of the MRCA of *D. versipellis s. lat.* from eastern (‘peripheral’) forms that later gave rise to *D. pleiantha*, was vicariant allo-peripatric divergence at the extreme edge of an ancestral species’ range, perhaps triggered by the onset of the Pleistocene climate oscillations (c. 2.4–1.8 Ma). In other words, the peripheral forms of *D. pleiantha* would have persisted in East China under deteriorating conditions but became isolated by the retracting range of adjacent populations. Notably, this scenario resembles the rarely recognized ‘centrifugal speciation model’ (CSM) of allo-peripatric speciation^{27–29}, and is also supported by the ENM results for the LGM (see below).

Our cpDNA-based estimates of demographic history (Table 2) provided clear evidence for spatial expansions in both the central-eastern clade of *D. versipellis s. lat.* during the mid-Pleistocene [c. 0.36 (0.20–0.61) Ma], possibly coinciding with the end of China’s ‘Penultimate Interglacial Period’ (c. 0.33–0.46 Ma^{30,31}), and in *D. pleiantha* during the Late Pleistocene/Holocene [c. 0.04 (0.0001–0.15) Ma]. Yet, apparently, these expansions induced no secondary contact, as both taxa are identified here as genetically distinct units in terms of cpDNA haplotype (Fig. 1a,c) and EST-SSR variation (Fig. 3). Moreover, the MIGRATE analyses of EST-SSRs uncovered extremely low levels of post-divergence gene flow between *D. versipellis* and *D. pleiantha* (Table 3), suggesting that their propensity for hybridization has been overstated in the taxonomic literature¹⁶. Rather, we suspect that both taxa, and their associated forest biomes, largely remained isolated from each other throughout the glacial and inter-/postglacial periods of the Quaternary. In fact, our ENM results indicate that, during the LGM (and possibly earlier cold periods), *D. versipellis s. lat.* and *D. pleiantha* occupied distinct refugia in West/Southwest China and East China, respectively, with only little potential range overlap in Central-East China (Fig. 4c,d). That glacial periods caused the range of *D. versipellis s. lat.* to retract south-westward (i.e. around the Sichuan Basin, Yungui Plateau), and/or to become fragmented, is similar with findings in other WTD forest species from subtropical China (e.g. *Cercidiphyllum japonicum*⁶; *Kalopanax septemlobus*⁵; but see below). However, that East China (and areas around the mid-/lower Yangtze) offered potentially suitable climate conditions for *D. pleiantha* during the LGM contrasts

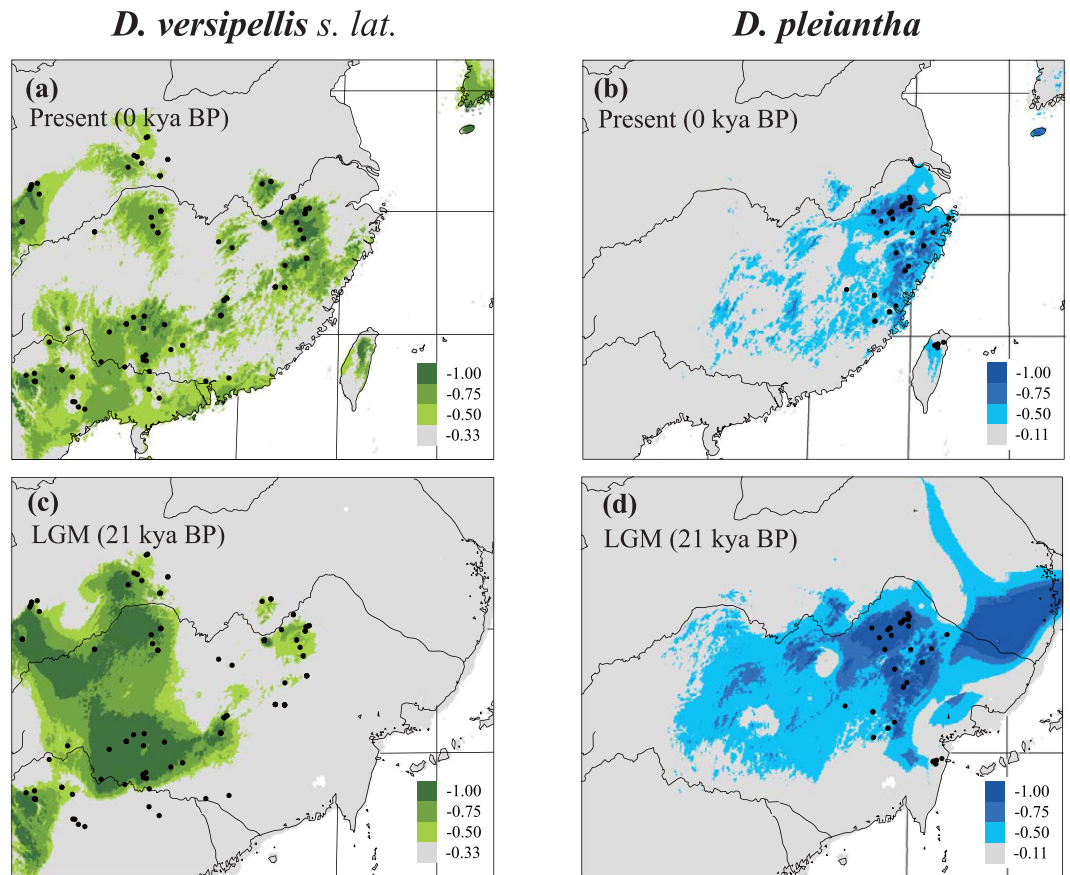


Figure 4. Predicted distributions of *D. versipellis s. lat.* (including *D. versipellis*, *D. difformis* and *D. majoensis*), and *D. pleiantha* under (a,b) current climate conditions (1950–2000) and (c,d) the Last Glacial Maximum (LGM; 21 kya BP) based on ecological niche modelling using MAXENT v3.3.1 (<http://www.cs.princeton.edu/~schapire/maxent/>). Maps were generated using ARCGIS v9.3 (ESRI, Redlands, CA, USA).

with fossil-pollen data^{3,32}, suggesting that WTE forest largely retracted to the tropical South ($\leq 24^{\circ}\text{N}$). Instead, our ENM data suggest that this drought-sensitive herb¹⁶ benefitted from relative climatic stability in East China over the last glacial cycle(s), possibly due to warm and moist air masses from the ocean^{33,34} and/or topographic heterogeneity (e.g. Huang/Tianmu Mts.) that buffered against regional climate change³⁵. The almost complete fixation of cpDNA haplotype H48 in *D. pleiantha* (Fig. 1a), together with its homogenous EST-SSR gene pool (red in Fig. 3), precludes more detailed inferences about the glacial refugia and (re-)colonization routes of this species. Nonetheless, when combined with the MDA evidence, this near lack of genetic structure most likely results from a founder effect in the wake of a leading edge expansion during the Late Pleistocene/Holocene⁶. Assuming this expansion signals recovery from a bottleneck^{36,37}, any such earlier reduction in population size in *D. pleiantha* could have further enhanced its geographic isolation from *D. versipellis s. lat.* Moreover, any climate amelioration should have likewise promoted the isolation of these taxa. This hypothesis is based on the expectation that inter-/postglacial periods generally fostered the displacement of WTD forest into higher elevations by the expansion of WTE forest at lower elevations, as supported by phylogeographic studies^{31,38} and palaeo-data³². In sum, these findings indicate that the divergence between *D. versipellis s. lat.* and *D. pleiantha* has been triggered and maintained by geographical and ecological-adaptive processes in response to Pleistocene climate change.

Our survey of cpDNA sequence variation at 31 locations of *D. versipellis s. lat.* revealed substantial population differentiation ($\Phi_{ST} = 0.84$) and significant phylogeographic structure ($N_{ST} = 0.818 > G_{ST} = 0.611$; $P < 0.05$). These patterns largely result from the major phylogeographic break across the Sichuan Basin between ancestral ‘western’ and derived (‘founded’) central-eastern populations, as well as the marked differences within the latter (Fig. 1). These results are similar to those of two earlier studies in *D. versipellis* based on cpDNA³¹ and AFLPs³⁹, which also registered a major west vs. central-east split in this species, despite more limited samples (10 and nine populations, respectively). By contrast, based on our EST-SSR data, STRUCTURE identified three geographical clusters in *D. versipellis s. lat.* (west-north, south, east). Hence, all populations of the western cpDNA lineage (EM, DJ) were found to share the same gene pool with five northern populations (ZF, SN, MHP, TP, TT) of the central-eastern lineage (Fig. 3a,b). This failure of the EST-SSRs to register the phylogeographic cpDNA split is probably best explained by extended times for lineage sorting to occur at nuclear compared to cytoplasmic loci⁴⁰ rather than contemporary pollen-mediated introgression from the north into the west. The latter appears less likely if one considers not only the limited pollen-dispersal capacity of *Dyosma*^{16,39} but also the MIGRATE results (Table 3), indicating only low levels of gene flow between the three EST-SSR clusters of *D. versipellis s. lat.* The previous

cpDNA analysis of *D. versipellis*³¹ dated the separation between the western and central-eastern lineages to the mid-Pleistocene, about 0.48 Ma (95% CI: 0.18–0.86 Ma). In the present study, ABC analyses revealed that the divergence of the three regional EST-SSR clusters of *D. versipellis s. lat.* likewise dates back to the mid-Pleistocene (c. 0.59 Ma, 95% CI: 0.15–1.29 Ma; t_2 in Supplementary Fig. S5a, Table 4). Even though difficult to compare across studies, and despite their broad confidence intervals, both time estimates strongly suggest that major lineage divergences within *D. versipellis s. lat.* long pre-date the LGM. Moreover, the present cpDNA and EST-SSR data provide a congruent picture of the phylogeographic history of *D. versipellis s. lat.*, as both support the hypothesis that populations of this taxon survived the (Late) Pleistocene glacial maxima in multiple refugia in both West/Southwest China and Central-East China. In the former region, there were probably three refugia: one, as identified by cpDNA, at the eastern slopes of the QTP (comprising populations DJ, EM; Fig. 1a); and two others, as revealed by the EST-SSRs, extending south vs. north-east of the Sichuan Basin (light vs. dark blue clusters; Fig. 3b,c). In Central-East China, there was probably a major refuge located south of the mid-/lower Yangtze, as demonstrated by the nuclear cohesion of all eastern populations (green cluster; Fig. 3b,c), and possibly a more localized one further north in the Dabie Mts., as might be inferred from the high number of unique cpDNA haplotypes in the ‘north-cluster’ population TT (Fig. 1a, Supplementary Table S2). Notably, these latter two refugia are also predicted by the ENM of *D. versipellis s. lat.* for the LGM (Fig. 4c); on the other hand, the genetic data better serve to illustrate that West/Southwest China did not function as a single refugium for this WTD forest species during glacial periods but instead hosted different nuclear and/or cytoplasmic lineages in separate refugia. The occurrence of these ‘refugia-within-refugia’⁴¹ not only caused the allopatric divergence of these lineages but probably also fostered the recent speciation of *D. difformis* and *D. majoensis* from within *D. versipellis*.

Dysosma difformis and *D. majoensis* have proximate but non-overlapping eastern vs. western distributions along the southern fringe of the Sichuan Basin (Fig. 1a,b). These two narrow endemics share the same WTD forest habitat with *D. versipellis (s. str.)* and are sometimes difficult to distinguish from the latter based on morphology, as most of their diagnosable features relate to aspects of leaf lobing^{16,18}. Our genetic data clearly suggest a close relationship among these three taxa, without traces of other *Dysosma* species (e.g. *D. delavayi*) as suggested previously for *D. majoensis*¹⁶. Rather, all cpDNA haplotypes unique to *D. difformis* (H52–H54; only population SH) and the single one fixed in *D. majoensis* (H55) are independent mutational derivatives of the most common and widespread haplotype of *D. versipellis* (H1; see Fig. 1a,c). Moreover, the EST-SSR data show that *D. difformis* and *D. majoensis*, despite sharing the same gene pool (light blue) with all southern populations of *D. versipellis* (Fig. 3), have likely exchanged only few migrants with each other and the latter species in the distant past (Table 3). Finally, the corresponding ABC analysis favoured a scenario of unresolved relationships, according to which *D. difformis*, *D. majoensis* and *D. versipellis* (southern cluster) evolved independently from a common ancestor, possibly during the mid-Pleistocene [c. 0.43 (0.07–1.47) Ma; t_4 in Supplementary Fig. S5b]. Taken together, these data seem to support a scenario in which climate deterioration following China’s penultimate interglacial period (c. 0.33–0.46 Ma; see above) promoted the independent origin of *D. difformis* and *D. majoensis* from within *D. versipellis* through population isolation in two separate refugia along the southern Sichuan Basin. If so, the presence of cpDNA haplotypes H1 and H55 in *D. difformis* (Fig. 1a,c) most likely results from, respectively, incomplete lineage sorting (ILS; i.e. recent descent from *D. versipellis*) and past chloroplast introgression from *D. majoensis*, perhaps facilitated by massive expansion of suitable habitat in the area during the LGM (Fig. 4c). Clearly, the present data cannot fully disentangle these hypotheses of recent speciation, ILS and/or past admixture; however, they do suggest that the combined effects of these historical processes, rather than contemporary gene flow, are responsible for the limited evidence of genetic (cytoplasmic/nuclear) divergence of *D. difformis* and *D. majoensis*. These data, therefore, should not be interpreted to reject the hypothesis of geographic-reproductive isolation of these two endemics and their evolutionary distinctiveness as taxonomic units of interest^{42,43}; rather, they underscore the mountain regions of Southwest China of central importance for (micro-)refugial allopatric speciation^{31,44}.

In summary, the data presented here for the *D. versipellis-pleiantha* complex suggest an overriding role of Quaternary climate change in triggering essentially allopatric (incipient) speciation in this group of forest-restricted plant species in subtropical China.

Materials and Methods

Plant material and sampling design. The ‘*D. versipellis-pleiantha*’ complex comprises four closely related species (*D. pleiantha*, *D. versipellis*, *D. difformis* and *D. majoensis*) of the diploid ($2n = 12$) perennial understory herb genus *Dysosma* Woodson (Berberidaceae, Podophylloideae; c. 7 spp.). All four species have self-incompatible, dull/red-coloured, and insect (carrion fly/beetle) pollinated flowers that mature into purplish-red berries^{16,39,45,46}, and, like all species of *Dysosma*, are able to produce fertile F_1 hybrids^{16,46}. In this study, we sampled a total of 618 individuals from 42 localities of *D. versipellis* (24), *D. pleiantha* (11), *D. difformis* (4), and *D. majoensis* (3) (Supplementary Table S2, Fig. 1). Together, this sampling covers most of the distribution range of this species complex. Voucher specimens representative of all populations sampled are stored at the Herbarium of Zhejiang University (HZU). Of these 42 populations, five populations were only surveyed for either cpDNA (pop. JF and NG) or EST-SSR loci (pop. MT, ZP and QL), given the small amount of leaf material available. Accordingly, a total of 577 individuals from 40 populations were employed for the EST-SSR analysis, whereas 452 individuals from 39 populations were also surveyed for cpDNA sequence variation (Supplementary Table S2). For the phylogenetic cpDNA analyses (see below), we also collected samples from the three remaining *Dysosma* species, i.e. *D. aurantiocaulis* (Handel-Mazzetti) Hu (1 individual), *D. delavayi* (6), and *D. tsayuensis* T.S. Ying (1), plus one individual each of the closely related *Sinopodophyllum hexandrum* (Royle) Ying and *Podophyllum peltatum* L.^{19,47}.

DNA extraction, cpDNA sequencing, and EST-SSR genotyping. Total genomic DNA was extracted from the silica-dried leaf materials using DNA Plantzol (Invitrogen) following the manufacturer’s instructions.

For the phylogeographic and phylogenetic cpDNA analyses, we sequenced three non-coding intergenic spacer (IGS) regions, i.e. *trnL-trnF*, *trnL-ndhJ*⁴⁸ and *trnS-trnM*⁴⁹. Sequences were generated with an ABI 377XL DNA sequencer (Applied Biosystems), and edited, assembled, and aligned in GENEIOUS v7.1.7⁵⁰.

Fifteen EST-SSR primer pairs, developed from the transcriptome of *D. versipellis* (termed 'EDV' primers), were employed for genotyping all 577 DNA samples. Ten of those primers (EDV-30, 37, 40, 46, 52, 53, 54, 59, 60, 67) were described in Guo *et al.*⁵¹ (GenBank accession numbers: KJ000290–KJ000299). Respective primers for the other five loci (EDV-81, 82, 102, 118, 119) were developed in this study (Supplementary Table S6). All loci were proved to be consistently variable and transferable among the four species of the *D. versipellis-pleiantha* complex after a pilot study. Amplification of the EST-SSR loci followed the protocol of Guo *et al.*⁵¹. Fluorescently labelled PCR products were supplemented with the internal size standard GS-500 and separated on a MegaBACE 1000 (GE Healthcare Biosciences, Sunnyvale, CA, USA). Alleles were scored manually in GENEMARKER v2.20 (SoftGenetics, State College, PA, USA).

Chloroplast haplotype diversity and population structure. The number of haplotypes, haplotype diversity (h), and nucleotide diversity (π) were calculated at the levels of populations (h_s , π_s) and species (h_T , π_T) using DNASP v5.0⁵². Non-hierarchical and hierarchical analyses of molecular variance (AMOVAs) were carried out in ARLEQUIN v3.5⁵³ with significance of Φ -statistics tested by 1000 non-parametric random permutations. The presence of phylogeographic structure was tested across all populations of the species complex and for each species and/or lineage separately using PERMUT v1.0⁵⁴. All haplotype sequences identified in the present study were deposited in GenBank with accession numbers (see Supplementary Table S7 for details).

Phylogenetic haplotype relationships and molecular dating. Phylogenetic cpDNA haplotype trees of the species complex were constructed using maximum likelihood (ML) and Bayesian inference (BI) methods, with the remaining *Dysosma* species treated as part of the ingroup and *S. hexandrum* and *P. peltatum* as outgroup (see Supplementary Methods for details). For the phylogeographic cpDNA analyses, a haplotype network was constructed in TCS v1.21⁵⁵ under the 95% statistical parsimony criterion with gaps (indels) coded as substitutions (A or T), and rooted with *D. delavayi*.

Divergence date analyses were conducted on the cpDNA data set in BEAST v1.80^{56,57} using the GTR + G substitution model (see Supplementary Methods for details) under the assumption of an uncorrelated lognormal relaxed clock⁵⁶. As there are no known fossil records in Podophylloideae, we employed a substitution rate of 1.3×10^{-9} substitutions per site per year (s/s/y) as estimated for non-coding chloroplast regions of *D. versipellis*³¹. We implemented two coalescent-type models, assuming either a constant population size or population expansion. For each model, four independent Markov chain Monte Carlo (MCMC) runs of 50 million generations each were performed with sampling every 5000 generations, following a burn-in of the initial 10% cycles. We used TREEANNOTATOR v1.6.1⁵⁷ to construct a maximum clade credibility tree.

Demographic analyses. To infer the demographic history of major cpDNA lineages within the species complex, we tested the null hypotheses of a spatial expansion and a pure demographic expansion using mismatch distribution analysis (MDA) in ARLEQUIN. For two expanding lineages identified, the MDA-derived expansion parameter (τ) and its 95% confidence interval (CI) were converted to absolute estimates of time since expansion (see Supplementary Methods for more details). In addition, we used tests of selective neutrality [Tajima's D ⁵⁸, Fu's F_s ⁵⁹] to infer potential population growth and expansion.

Population structure and migration/gene flow inferred from EST-SSR markers. For the EST-SSR dataset, all 15 genotyped loci were checked for frequencies of null alleles, deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) (see Supplementary Methods for details). For each locus, we also tested potential signatures of selection in LOSITAN⁶⁰, ARLEQUIN, and BAYESCAN v2.1⁶¹. Following De Mita *et al.*⁶², we considered a locus to be under selection if at least two outlier tests were significant for that particular locus. Using FSTAT, the following diversity and inbreeding parameters were calculated for each population and across all 15 loci: N_A , numbers of alleles; R_s , allelic richness⁶³; H_E ; and the average inbreeding coefficient (F_{IS}). Hierarchical and non-hierarchical AMOVAs were carried out in ARLEQUIN using R -statistics and significance tests as described above for cpDNA.

Genetic subgroups of the species complex were identified by Bayesian analyses in STRUCTURE v2.3.4⁶⁴, using the admixture model and assuming independent allele frequencies among populations (see Supplementary Methods for details). Clustering analyses for the EST-SSRs were complemented with an individual-based principal coordinates analysis (PCoA) using software GENALEX v6.5⁶⁵.

We also used coalescent-based methods implemented in MIGRATE-N v3.6⁶⁶ to obtain pairwise estimates of mutation-scaled migration rate (M) over $c. 4N_e$ generations in the past⁶⁷ between species and/or intraspecific regional clusters (only *D. versipellis*; see STRUCTURE results). All analyses were carried out with and without outlier loci included to assess their effects on the inference of population structure and gene flow⁶⁸. More software running details can be found in the Supplementary Materials.

Testing for the times and orders of species/cluster divergence using ABC. We used ABC simulations in DIYABC v2.0^{69,70} to gain further insights into the times and orders of divergence among clusters identified by STRUCTURE based on the nine neutral EST-SSR loci. In a first analysis ('a'), we tested the simultaneous divergence of the three clusters of '*D. versipellis* s. lat.' (i.e. west-north, east, south, with the latter including *D. difformis*/*D. majoensis*; see Results) from a common ancestor against three alternative models, reflecting all possible relationships among these clusters (see scenarios 1–4 in Supplementary Fig. S5a and Table S5). For each scenario, *D. pleiantha* was assumed as sister, based on the cpDNA haplotype network analysis (see Results). In

a second analysis ('b'), we tested four similar divergence order hypotheses between the taxonomic units of the southern cluster, i.e. *D. versipellis* (south), *D. difformis* and *D. majoensis* (scenarios 1–4 in Supplementary Fig. S5b and Table S5). Additional details in data simulation, scenario comparison, parameter estimation, etc. can be found in the Supplementary Materials.

Ecological niche modelling and niche identity tests. Ecological niche modelling (ENM) was carried out in MAXENT v3.3.1^{71,72} to predict suitable past and present climate envelopes for, respectively, the WTD forest species (i.e. *D. versipellis* s. lat., including *D. versipellis*, *D. difformis*, *D. majoensis*) and the WTE forest species *D. pleiantha*. Information on the geographic distribution of the three former species was based on the 31 populations included in this study, combined with 102 presence records obtained from the Chinese Virtual Herbarium (CVH) (www.cvh.org.cn) and the National Specimen Information Infrastructure of China (NSII) (www.nsi.org.cn). In addition, 44 records were available for *D. pleiantha*, i.e. 11 localities of this study plus 33 obtained from CVH and NSII.

Current distributions of the two sets of species were modelled using six bioclimatic data layers (annual mean temperature, annual precipitation, precipitation of wettest, driest, warmest and coldest quarter) available from the WorldClim database (<http://www.worldclim.org>)⁷³ at 2.5 arc-min resolution for the present (1950–2000). This restricted bioclimatic dataset avoided including highly correlated variables, and thus prevented potential over-fitting⁷⁴. This model was then projected onto the palaeoclimate dataset simulated by the COMMUNITY CLIMATE SYSTEM MODEL v3.0⁷⁵ to infer the extent of suitable habitat during the Last Glacial Maximum (LGM; c. 21 kya BP). Model performance was evaluated using the area under the 'Receiver Operating Characteristic (ROC) Curve' (AUC) calculated by MAXENT. Values between 0.7 and 0.9 indicate good discrimination⁷⁶.

We also performed niche identity tests in ENMTOOLS v1.4.3⁷⁷ based on all the 19 BIOCLIM variables from the WorldClim dataset for testing the null hypothesis that the WTD vs. WTE forest species (as defined above) are occupying identical climatic environments ('niches'). However, we did not further evaluate the niche similarity between *D. difformis* and *D. majoensis* because of the very low number of known occurrence records for either species. Niche overlap was quantified using the standardized Hellinger distance (*I*) and Schoener's *D*⁷⁸.

Data Availability. cpDNA (trnL–trnF, trnF–ndhJ and trnS–trnfM) sequences obtained in this study are deposited in GenBank (see Supplementary Table S7 for accession numbers). Microsatellite genotypes and maxent input file have been archived with Dryad (doi: 10.5061/dryad.9d8k4).

References

1. Qian, H. & Ricklefs, R. E. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* **407**, 180–182 (2000).
2. Hou, H. & Xue-Yu, H. Vegetation of China with reference to its geographical distribution. *Ann. Mo. Bot. Gard.* **3**, 509–549 (1983).
3. Yu, G. *et al.* Palaeovegetation of China: a pollen data-based synthesis for the mid-Holocene and last glacial maximum. *J. Biogeogr.* **27**, 635–664 (2000).
4. Qiu, Y. X. *et al.* Molecular phylogeography of East Asian *Kirengeshoma* (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. *New Phytol.* **183**, 480–495 (2009).
5. Sakaguchi, S. *et al.* Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree *Kalopanax septemlobus* (Araliaceae) in East Asia. *Mol. Ecol.* **21**, 3823–3838 (2012).
6. Qi, X. S. *et al.* Molecular data and ecological niche modelling reveal a highly dynamic evolutionary history of the East Asian Tertiary relict *Cercidiphyllum* (Cercidiphyllaceae). *New Phytol.* **196**, 617–630 (2012).
7. Bai, W. N., Wang, W. T. & Zhang, D. Y. Phylogeographic breaks within Asian butternuts indicate the existence of a phylogeographic divide in East Asia. *New Phytol.* **209**, 1757–1772 (2016).
8. Zhang, Y., Wang, I. J., Comes, H. P., Peng, H. & Qiu, Y. Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure in a Tertiary relict species, *Emmenopterys henryi* (Rubiaceae). *Sci. Rep.* **6**, 24041 (2016).
9. Cao, Y., Comes, H. P., Sakaguchi, S., Chen, L. & Qiu, Y. Evolution of East Asia's Arcto-Tertiary relict *Euptelea* (Eupteleaceae) shaped by Late Neogene vicariance and Quaternary climate change. *BMC Evol. Biol.* **16**, 1 (2016).
10. Gao, L. M. *et al.* High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Mol. Ecol.* **16**, 4684–4698 (2007).
11. Li, Y., Yan, H. F. & Ge, X. J. Phylogeographic analysis and environmental niche modeling of widespread shrub *Rhododendron simsii* in China reveals multiple glacial refugia during the last glacial maximum. *J. Syst. Evol.* **50**, 362–373 (2012).
12. Sun, Y., Hu, H., Huang, H. & Vargas-Mendoza, C. F. Chloroplast diversity and population differentiation of *Castanopsis fargesii* (Fagaceae): a dominant tree species in evergreen broad-leaved forest of subtropical China. *Tree Genet. Genomes* **10**, 1531–1539 (2014).
13. Wang, Y. H. *et al.* 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.* **206**, 852–867 (2015).
14. Qiu, Y., Fu, C. & Comes, H. P. 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.* **59**, 225–244 (2011).
15. Liu, J. Q., Sun, Y. S., Ge, X. J., Gao, L. M. & Qiu, Y. X. Phylogeographic studies of plants in China: advances in the past and directions in the future. *J. Syst. Evol.* **50**, 267–275 (2012).
16. Shaw, J. M. The genus *Podophyllum* in *The genus Epimedium and other herbaceous Berberidaceae* (eds Green, P. S. & Mathew, B.) 239–314 (Timber Press, Portland, OR, 2002).
17. Mao, Y. R., Zhang, Y. H., Nakamura, K., Guan, B. C. & Qiu, Y. X. Developing DNA barcodes for species identification in Podophylloideae (Berberidaceae). *J. Syst. Evol.* **52**, 487–499 (2014).
18. Ying, T. S., Boufford, D. E. & Brach, A. R. Berberidaceae in *Flora of China* (eds Wu, Z. Y., Raven, P. H.) **19**, 783–787 (Science Press, Beijing, China & Missouri Botanical Garden Press, St. Louis, USA, 2011).
19. Wang, W., Chen, Z., Liu, Y., Li, R. & Li, J. Phylogenetic and biogeographic diversification of Berberidaceae in the Northern Hemisphere. *Syst. Bot.* **32**, 731–742 (2007).
20. Gamisch, A., Fischer, G. A. & Comes, H. P. Frequent but asymmetric niche shifts in *Bulbophyllum* orchids support environmental and climatic instability in Madagascar over Quaternary time scales. *BMC Evol. Biol.* **16**, 1 (2016).

21. Liu, Y., Wang, Y. & Huang, H. Species-level phylogeographical history of *Myricaria* plants in the mountain ranges of western China and the origin of *M. laxiflora* in the Three Gorges mountain region. *Mol. Ecol.* **18**, 2700–2712 (2009).
22. Knox, E. B. & Palmer, J. D. The origin of *Dendrosenecio* within the Senecioneae (Asteraceae) based on chloroplast DNA evidence. *Am. J. Bot.* **82**, 1567–1573 (1995).
23. Nürk, N. M., Madriñán, S., Carine, M. A., Chase, M. W. & Blattner, F. R. Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). *Mol. Phylogenet. Evol.* **66**, 1–16 (2013).
24. Qiang, X. K., Li, Z. X., Powell, C. M. & Zheng, H. B. Magnetostratigraphic record of the Late Miocene onset of the East Asian monsoon, and Pliocene uplift of northern Tibet. *Earth Planet. Sci. Lett.* **187**, 83–93 (2001).
25. Clift, P. D. *et al.* Correlation of Himalayan exhumation rates and Asian monsoon intensity. *Nat. Geosci.* **1**, 875–880 (2008).
26. López-Antoñanzas, R., Knoll, F., Wan, S. & Flynn, L. J. Causal evidence between monsoon and evolution of rhizomyine rodents. *Sci. Rep.* **5**, 9008 (2015).
27. Brown, W. L. Centrifugal speciation. *Q. Rev. Biol.* **32**, 247–277 (1957).
28. Frey, J. K. Modes of peripheral isolate formation and speciation. *Syst. Biol.* **42**, 373–381 (1993).
29. Briggs, J. C. Centrifugal speciation and centres of origin. *J. Biogeogr.* **27**, 1183–1188 (2000).
30. Yi, C., Cui, Z. & Xiong, H. Numerical periods of Quaternary glaciations in China. *Quatern. Sci.* **25**, 609–619 (2005).
31. Qiu, Y., Guan, B., Fu, C. & Comes, H. P. Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. *Mol. Phylogenet. Evol.* **51**, 281–293 (2009).
32. Harrison, S. P., Yu, G., Takahara, H. & Prentice, I. C. Palaeovegetation (Communications arising): diversity of temperate plants in east Asia. *Nature* **413**, 129–130 (2001).
33. Zhang, R. & Liu, X. D. The effects of tectonic uplift on the evolution of Asian summer monsoon climate since Pliocene. *Chinese J. Geophys. CH.* **53**, 948–960 (2010).
34. Ao, H. *et al.* Different orbital rhythms in the Asian summer monsoon records from North and South China during the Pleistocene. *Glob. Planet. Change* **80**, 51–60 (2012).
35. Gavin, D. G. *et al.* Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytol.* **204**, 37–54 (2014).
36. Rogers, A. R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569 (1992).
37. Arenas, M., Ray, N., Currat, M. & Excoffier, L. Consequences of range contractions and range shifts on molecular diversity. *Mol. Biol. Evol.* **29**, 207–218 (2012).
38. Qi, X., Yuan, N., Comes, H. P., Sakaguchi, S. & Qiu, Y. A strong 'filter' effect of the East China Sea land bridge for East Asia's temperate plant species: inferences from molecular phylogeography and ecological niche modelling of *Platycrater arguta* (Hydrangeaceae). *BMC Evol. Biol.* **14**, 1 (2014).
39. Guan, B., Fu, C., Qiu, Y., Zhou, S. & Comes, H. P. Genetic structure and breeding system of a rare understory herb, *Dysosma versipellis* (Berberidaceae), from temperate deciduous forests in China. *Am. J. Bot.* **97**, 111–122 (2010).
40. Joly, S., McLenachan, P. A. & Lockhart, P. J. A statistical approach for distinguishing hybridization and incomplete lineage sorting. *Am. Nat.* **174**, E54–E70 (2009).
41. Gómez, A. & Lunt, D. H. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula in *Phylogeography of Southern European Refugia*. (eds Weiss, S., Ferrand, N.) Ch 5, 155–188 (Springer Press, Dordrecht, The Netherlands, 2007).
42. Moritz, C. Defining 'evolutionarily significant units' for conservation. *Trends. Ecol. Evol.* **9**, 373–374 (1994).
43. Wang, J. Y., Frasier, T. R., Yang, S. C. & White, B. N. Detecting recent speciation events: the case of the finless porpoise (genus *Neophocaena*). *Heredity* **101**, 145–155 (2008).
44. Liu, J. *et al.* Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytol.* **199**, 1093–1108 (2013).
45. Shaw, J. New combinations for the varieties of *Sinopodophyllum hexandrum*. *Hanburyana* **4**, 33–39 (2009).
46. Gong, X., Guan, B., Zhou, S. & Ge, G. Reproductive biology of the rare plant, *Dysosma pleiantha* (Berberidaceae): breeding system, pollination and implications for conservation. *Pak. J. Bot.* **47**, 951–957 (2015).
47. Xiang, Q., Soltis, D. E., Soltis, P. S., Manchester, S. R. & Crawford, D. J. Timing the eastern Asian–eastern North American floristic disjunction: molecular clock corroborates paleontological estimates. *Mol. Phylogenet. Evol.* **15**, 462–472 (2000).
48. Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* **17**, 1105–1109 (1991).
49. Demesure, B., Sodzi, N. & Petit, R. J. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Mol. Ecol.* **4**, 129–134 (1995).
50. Drummond, A. J. *et al.* Geneious v4.8.5. <http://www.geneious.com> (2009).
51. Guo, R. *et al.* Characterization and cross-species transferability of EST–SSR markers developed from the transcriptome of *Dysosma versipellis* (Berberidaceae) and their application to population genetic studies. *Mol. Breeding* **34**, 1733–1746 (2014).
52. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
53. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* **10**, 564–567 (2010).
54. Pons, O. & Petit, R. J. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**, 1237–1245 (1996).
55. Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1659 (2000).
56. Drummond, A. J., Nicholls, G. K., Rodrigo, A. G. & Solomon, W. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**, 1307–1320 (2002).
57. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).
58. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
59. Fu, Y. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925 (1997).
60. Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A. & Luikart, G. LOSITAN: a workbench to detect molecular adaptation based on a F_{ST} -outlier method. *BMC Bioinform.* **9**, 1 (2008).
61. Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–993 (2008).
62. Mita, S. *et al.* Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol. Ecol.* **22**, 1383–1399 (2013).
63. El Mousadik, A. & Petit, R. J. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theor. Appl. Genet.* **92**, 832–839 (1996).
64. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).

65. Peakall, P. E. & Smouse, R. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537–2539 (2012).
66. Beerli, P. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22**, 341–345 (2006).
67. Beerli, P. & Felsenstein, J. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* **98**, 4563–4568 (2001).
68. Soto-Cerda, B. J. & Cloutier, S. Outlier loci and selection signatures of simple sequence repeats (SSRs) in flax (*Linum usitatissimum* L.). *Plant Mol. Biol. Report* **31**, 978–990 (2013).
69. Cornuet, J. *et al.* Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713–2719 (2008).
70. Cornuet, J. *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* **30**, 1187–1189 (2014).
71. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model* **190**, 231–259 (2006).
72. Phillips, S. J. & Dudík, M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161–175 (2008).
73. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
74. Peterson, A. T. & Nakazawa, Y. Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Glob. Ecol. Biogeogr.* **17**, 135–144 (2008).
75. Collins, W. D. *et al.* The community climate system model version 3 (CCSM3). *J. Climate* **19**, 2122–2143 (2006).
76. Swets, J. A. Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–1293 (1988).
77. Warren, D. L., Glor, R. E. & Turelli, M. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607–611 (2010).
78. Schoener, T. W. The Anolis lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49**, 704–726 (1968).
79. Harpending, H. C. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* **66**, 591–600 (1994).
80. Schneider, S. & Excoffier, L. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152**, 1079–1089 (1999).

Acknowledgements

We thank Dr. Shota Sakaguchi (Kyoto University, Japan) and Dr. Feng Li (Northwest University, China) for help with, respectively, the ENM and ABC simulation analyses, and Prof. Feng Sheng Hu (University of Illinois, USA) for very insightful comments that improved earlier drafts of the manuscript. This research was supported by the National Natural Science Foundation of China (Grant Nos. 31170200, 31370241 and 31570214), Biodiversity International Collaborative Project of NSFC-NSF (No. 31461123001), and the Zhejiang Provincial Funds for Distinguished Young Scientists (Grant No. LR12C02001).

Author Contributions

Y.X.Q. conceived and designed the study. Y.R.M. and R.G. collected samples and genetic data. Y.H.W., Y.R.M. and Y.N.C. analyzed data. Y.H.W. and Y.R.M. wrote previous drafts of the manuscript. Y.X.Q. and H.P.C. led the writing of the final manuscript with help from all other authors.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wang, Y.-H. *et al.* Quaternary climate change drives allo-peripatric speciation and refugial divergence in the *Dysosma versipellis-pleiantha* complex from different forest types in China. *Sci. Rep.* **7**, 40261; doi: 10.1038/srep40261 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017