Genome wide sequencing provides evidence of adaptation to heterogeneous environments for the ancient relictual *Circaeaster agrestis* (Circaeasteraceae, Ranunculales)

Running title: Adaptive evolution of *Circaeaster agrestis*

Xu Zhang^{1,2,3,#}, Yanxia Sun^{1,2,#,*}, Jacob B. Landis^{4,5}, Jianwen Zhang⁶, Linsen Yang⁷, Nan Lin^{1,2,3}, Huajie Zhang^{1,2}, Rui Guo^{1,2,3}, Lijuan Li^{1,2,3}, Yonghong Zhang⁶, Tao Deng⁶, Hang Sun^{6,*}, Hengchang Wang^{1,2,*}

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

²Center of Conservation Biology, Core Botanical Gardens, Chinese Academy of Sciences, Wuhan 430074, Hubei, China;

³University of Chinese Academy of Sciences, Beijing, 100049 China;

⁴Department of Botany and Plant Sciences, University of California Riverside, Riverside, CA, USA;

⁵School of Integrative Plant Science, Section of Plant Biology and the L.H. Bailey Hortorium, Cornell University, Ithaca, NY USA;

⁶Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, 650201 China;

⁷Hubei Key Laboratory of Shennongjia Golden Monkey Conservation Biology, Administration of Shennongjia National Park, Shennongjia, Hubei, China.

[#]These authors contributed equally to this work.

*Correspondence: sunyanxia@wbgcas.cn; sunhang@mail.kib.ac.cn; hcwang@wbgcas.cn

27 **Summary**

28 ● Investigating the interaction between environmental heterogeneity and local
29 adaptation is critical to understand the evolutionary history of a species, providing the
30 premise for studying the response of organisms to rapid climate change. However, for
31 most species how exactly the spatial heterogeneity promotes population divergence and
32 how genomic variations contribute to adaptive evolution remain poorly understood.

33 ● We examine the contributions of geographical and environmental variables to
34 population divergence of the relictual, alpine herb *Circaeaster agrestis*, as well as
35 genetic basis of local adaptation using RAD-seq and plastome data.

36 ● We detected significant genetic structure with an extraordinary disequilibrium of
37 genetic diversity among regions, and signals of isolation-by-distance along with
38 isolation-by-resistance. The populations were estimated to begin diverging in the late
39 Miocene, along with a possible ancestral distribution of the Hengduan Mountains and
40 adjacent regions. Both environmental gradient and redundancy analyses revealed
41 significant association between genetic variation and temperature variables. Genome-
42 environment association analyses identified 16 putatively adaptive loci related to
43 biotic and abiotic stress resistance.

44 ● Our genome wide data provide new insights into the important role of
45 environmental heterogeneity in shaping genetic structure, and access the footprints of
46 local adaptation in an ancient relictual species, informing conservation efforts.

47 **Keywords:** local adaptation, *Circaeaster agrestis*, RAD sequencing, genetic structure,
48 environmental heterogeneity, temperature variables, conservation

49

50 **Introduction**

51 A primary goal of phylogeography is understanding how geographical and ecological
 52 factors shape the genetic structure and evolutionary history of species. Environmental
 53 factors can act in concert and vary in space, resulting in divergent natural selection
 54 leading to genetic divergence during adaptation to heterogeneous environments
 55 (Savolainen *et al.*, 2007). Such genetic divergence along environmental gradients or
 56 across varied ecological habitats, can be indicative of local adaptation, a mechanism
 57 beneficial to the long-term persistence of populations and intensifying genetic
 58 differentiation among populations (Kawecki & Ebert, 2004; Conover *et al.*, 2009;
 59 Colautti & Barrett, 2013; Savolainen *et al.*, 2013; Lowry *et al.*, 2019). Understanding
 60 the genetic basis of local adaptation will provide insights into the response of organisms
 61 to ongoing climate change, aiding in conservation efforts (Aitken *et al.*, 2008; Ahrens *et al.*, 2019).

63 Divergent natural selection is the driving force of genetic divergence (Darwin,
 64 1859; Coyne & Orr, 2004), while gene flow can buffer genetic differentiation (Spieth,
 65 1974; Goicoechea *et al.*, 2019). Evolution resulting from local adaptation depends on
 66 both the strength of local selective pressures and the homogenizing effect of
 67 connectivity among different localities. Historically, isolation by distance (IBD)
 68 (Wright 1946) have been widely employed in studies of genetic differentiation among
 69 populations. Mcrae (2006) proposed the isolation-by-resistance (IBR) model, which
 70 predicts genetic connectivity among populations in a complex landscape and provides a
 71 flexible and efficient tool to account for spatial heterogeneity, improving the
 72 understanding of environmental effect on genetic structuring. Genetic drift may
 73 reinforce the establishment and maintenance of genetic differentiation, particularly in
 74 populations with small effective size (Wright, 1946; Kimura & Crow, 1964).
 75 Quantifying the relative contributions among natural selection, genetic drift and
 76 population connectivity to genetic divergence is crucial for understanding the role of
 77 local adaptation in adaptive evolution of a species (Savolainen *et al.*, 2007; Friis *et al.*,

78 2018).

79 Accessing the molecular basis of local adaptation and identifying selective drivers
80 is still challenging for species with limited genomic resources (Mayol *et al.*, 2019).
81 Nonetheless, with advancements in sequencing, it is now possible to study thousands of
82 loci improving our understanding of genome-wide effects of accumulating genetic
83 divergence and genomic properties that influence the process of local adaptation
84 (McCormack *et al.*, 2013; Ellegren, 2014; Seehausen *et al.*, 2014; Weigel & Nordborg,
85 2015). In particular, highly divergent loci identified by Genome-Environment
86 Association (GEA) analyses can be interpreted as potential targets of divergent
87 selection associated with population-specific environmental covariables (Coop *et al.*,
88 2010; Rellstab *et al.*, 2015; Hoban *et al.*, 2016; Forester *et al.*, 2018). However, for
89 most species, such as relictual alpine herbs, how exactly the spatial heterogeneity
90 promotes population diversification and how genomic variations contribute to adaptive
91 evolution, in particular their interaction, remains poorly understood.

92 In the present study, we focus on investigating the evolutionary history and local
93 adaptation of *Circaeaster agrestis* Maxim, an annual alpine herb with a relatively
94 narrow habitat (Fu & Bartholomew, 2001; Sun *et al.*, 2017). *Circaeaster* Maxim., with
95 its only alliance *Kingdonia* Balf.f. & W.W. Smith, constitute the early-diverging
96 eudicot family Circaeasteraceae (Ranunculales) (The Angiosperm Phylogeny et al.,
97 2016). Due to special morphological characters such as open dichotomous leaf
98 venation, this genus is of great interest to many botanists (Foster, 1971; Ren *et al.*,
99 2003). *Circaeaster* exhibits low morphological diversity with only one species,
100 garnering a status of critical endangerment (Wild Plants Under State Protection in
101 China). The distributions of *C. agrestis* is confined to the Qinghai-Tibetan Plateau
102 (QTP) and adjacent areas which exhibit a wide range of elevations from 2,100-5,000 m.
103 The uplift of the QTP has created extraordinarily geomorphological and climatic
104 diversity (Mulch & Chamberlain, 2006). Due to the high mountain barrier formed by
105 the Himalayas in the south, the central and western QTP (i.e. the Himalayas) are

106 characterized by a cold and dry climate. In contrast, the eastern QTP (i.e. the Hengduan
107 Mountains) and adjacent areas are associated with deep valleys and characterized
108 mainly by a warm and wet climate (Song *et al.*, 2010; Tang *et al.*, 2013; Lu & Guo,
109 2014; Favre *et al.*, 2015). Given the high degree of geographical and ecological
110 heterogeneity, the QTP comprises a promising model system for studying genetic
111 signatures of local adaptation.

112 Mountainous regions like the QTP are often centers of endemism and species
113 diversity hotspots (Myers *et al.*, 2000; Noroozi *et al.*, 2018) due to uplift-driven
114 diversification (Hughes & Atchison, 2015; Chen *et al.*, 2019). The uplift of the QTP
115 produced diverse habitats facilitating rapid population divergence, but also created
116 geographical barriers leading to isolation and fragmentation of populations (He & Jiang,
117 2014; Deng *et al.*, 2020), the main drivers of species extinction (Hughes, 2017).
118 Therefore, investigating genetic structure and local adaptation of species with a
119 relatively narrow habitat will act as a proxy to comprehensively understand the role of
120 geographical and environmental heterogeneity, while also serving as a pioneer in
121 addressing the issues of small patchy diversity in biodiversity hotspots, aiding
122 conservation and management efforts for threatened species.

123 Specifically, we generated and analyzed two types of genomic datasets: single
124 nucleotide polymorphisms (SNPs) derived from the restriction site associated DNA
125 sequencing (RAD-Seq) of 18 *C. agrestis* populations and 20 plastome sequences of *C.*
126 *agrestis* populations with other Ranunculales representatives. We hypothesized that
127 environmental heterogeneity may exert strong selective pressures on *C. agrestis*,
128 driving population divergence and generating genetic variation. The aim of the present
129 study is to illustrate the evolutionary patterns of population divergence interacting with
130 environmental variables and access the potential genetic basis of local adaptation. Here,
131 we raise the following questions: 1) how alpine environments affect genetic structure
132 within species and drive population divergence, and 2) how genetic variation is
133 associated with local adaptation? Answering these questions will help us understand

species adaptation to rapidly changing climate and help to develop sound conservation programs.

Materials and Methods

Sampling, library preparation and sequencing

A total of 139 individuals of *C. agrestis* from 18 localities were sampled (Table S1). Sampling locations were chosen based on existing occurrence records covering the geographic and climatic distribution of the species (Fig. S1). Our field collection followed the ethics and legality of the local government. All voucher specimens were deposited in the Herbarium of Wuhan Botanical Garden (HIB) (Table S1). Total DNA was extracted from silica gel-dried leaves with a modified CTAB (Cetyl trimethylammonium bromide) method (Yang *et al.*, 2014). For RAD library construction and sequencing, genomic DNA was digested with the restriction enzyme EcoRI in a 30 ul reaction followed by ligation of P1 adapter by T4 ligase. Fragments were pooled, randomly sheared, and size-selected to 350–550 bp. A second adapter (P2) was then ligated. The ligation products were purified and PCR amplified, followed by gel purification, and size selection to 350–550 bp. Agilent 2100 Bioanalyzer and qPCR were used to qualify and quantify the library. Paired-end sequencing was performed on two lanes of Illumina HiSeq 2000 at BGI-Shenzhen (Shenzhen, Guangdong, China).

For plastome sequencing, 20 samples representing all 18 localities were chosen. For each sample, a 500-bp DNA TruSeq Illumina (Illumina Inc., San Diego, CA, USA) sequencing library was constructed using 2.5–5.0 ng sonicated DNA as input. The libraries were quantified using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and real-time quantitative PCR. Libraries were multiplexed and sequenced using a 2×125 bp run on one lane of Illumina HiSeq 2000 at BGI-Shenzhen (Shenzhen, Guangdong, China).

162 **Processing of Illumina data**

163 For RAD-sequencing, Illumina reads were processed into RAD-tags using the
 164 STACKS v.2.30 software pipeline (Catchen *et al.*, 2013). Samples were initially
 165 demultiplexed and filtered with PROCESS_RADTAGS. Reads with an average Phred
 166 score of at least 30, an unambiguous barcode, and restriction cut site were retained. The
 167 Perl wrapper DENOVO_MAP.PL was used for executing USTACKS, CSTACKS, and
 168 SSTACKS (Catchen *et al.*, 2013). Following the suggested protocol (Rochette &
 169 Catchen, 2017), we investigated a range of parameter values with M and n values
 170 ranging from 1 to 9 (fixing $M = n$) and $m = 3$. We plotted the number of polymorphic
 171 loci shared across samples (the r80 loci) and the distribution of the number of SNPs per
 172 locus.

173 The POPULATIONS module in the STACKS pipeline was used to produce
 174 datasets for downstream population genetic analyses. Polymorphic RAD loci that were
 175 present in all 139 individuals were retained. To validate the influence of missing data,
 176 we also employed a filtering parameter of $-r = 0.8$, in which at least 80% of individuals
 177 in a population were required to process a locus. Potential homologs were excluded by
 178 removing loci showing heterozygosity > 0.5 . We further filtered our dataset with a
 179 minor allele frequency (MAF) > 0.01 and kept only biallelic SNPs. Finally, we filtered
 180 our dataset by selecting the most informative SNP based on the number of minor
 181 alleles.

182 For genome-skimming sequencing, raw sequence reads were filtered using
 183 Trimmomatic v.0.36 (Bolger *et al.*, 2014) by removing duplicate reads and
 184 adapter-contaminated reads. Remaining reads were directly mapped to the plastome of
 185 *C. agrestis* (NCBI accession number: KY908400; plastome size: 151,033 bp) using
 186 Geneious v9.0.2 (Kearse *et al.*, 2012). After mapping, the mean depth of coverage was
 187 325.61x (median: 323.55x), ranging from 91.81x (WLG) to 804.08x (ZF). Initial
 188 annotations were implemented in the Plastid Genome Annotator (PGA) (Qu *et al.*,
 189 2019), and then refined by manual correction in Geneious. The annotated plastomes

were deposited in GenBank (accession numbers MT228704-MT228722).

Population structure and genetic diversity

Population genetic structure was estimated using Bayesian clustering and principal coordinate analysis (PCoA). Bayesian clustering was performed in STRUCTURE v2.3.4 (Pritchard *et al.*, 2000) with the admixture model. As STRUCTURE analysis assumes that loci are unlinked, we thus filtered out linked SNPs using the `–write_single_snp` option in POPULATIONS (Catchen *et al.*, 2013), which exports only the first SNP per locus for analysis. To determine the optimal number of groups (K), we ran STRUCTURE 10 times for $K = 1$ to $K = 10$. Each run was performed for 200,000 Markov Chain Monte Carlo (MCMC) generations with a burn-in period of 100,000 generations. The optimal K was chosen using both delta- K method implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012), and cross-entropy criterion implemented in LEA v2.8.0 (*snmf* function) (<https://bioconductor.org/packages/LEA/>) (Frichot. & François, 2015). The *snmf* function estimates an entropy criterion that evaluates the quality of fit of the statistical model to the data using a cross-validation technique to choose the number of ancestral populations that best explains the genotypic data (Frichot *et al.*, 2014). The coefficient for cluster membership of each individual was averaged across the 10 independent runs using CLUMPP (Jakobsson & Rosenberg, 2007) and plotted using DISTRUCT (Rosenberg, 2004). We performed PCoA analyses using *dartR* v1.1.11 (*gl.pcoa* function). We tested the overall population structure using the G-statistic test (Goudet *et al.*, 1996) by *gstat.randtest* function in *adeigenet* v2.1.1 (Jombart, 2008). An analysis of molecular variance (AMOVA) was performed to quantify genetic differentiation among and within populations and genetic groups using Arlequin v3.5 (Excoffier *et al.*, 2007) with significance tests of variance based on 10,000 permutations.

We estimated genetic diversity indices including nucleotide diversity (π), expected heterozygosity (H_e), observed heterozygosity (H_o) and inbreeding coefficients (F_{IS})

218 using POPULATIONS. The polymorphic sites were exported in *Phylip* format and used
 219 for calculating haplotypes in DnaSP v6.12 (Rozas *et al.*, 2017). Population and group
 220 level F_{ST} using Weir and Cockerham's method (Weir & Cockerham, 1984) was done
 221 using *hierfstat* (Goudet, 2005) in R v3.6.1 (R Team, 2014). The absolute differentiation
 222 (D_{XY}) among genetic groups were measured using a Perl script provided by Ru *et al.*,
 223 (2018) (Li *et al.*, 2020). Finally, we assessed the contemporary effective population
 224 sizes (N_e) for each population and group using NeEstimator v2.1 (Do *et al.*, 2014).
 225 Estimates of N_e were calculated from the bias corrected (Waples, 2006) linkage
 226 disequilibrium method (Hill, 1981) with a minor allele frequency cutoff of 0.05 and 95%
 227 confidence intervals (CI) estimated by jackknifing.

228

229 **Population divergence**

230 We adopted a two-step strategy for investigating the divergent time of *C. agrestis*. First,
 231 we used plastome sequences assembled from Ranunculales and newly sampled *C.*
 232 *agrestis* populations to estimate the diversification of *C. agrestis*. A total of 47
 233 plastomes were included. The 79 protein coding regions (CDS) were extracted and
 234 aligned using MAFFT v7.313 (Kato & Standley, 2013) and then concatenated using
 235 PhyloSuite v1.1.15 (Zhang *et al.*, 2019). The program BEAST2 v2.5.2 (Bouckaert *et al.*,
 236 2014) was employed for molecular dating with concatenated matrix using a GTR+G+I
 237 substitution model selected by jModelTest v2.0.1 (Darriba *et al.*, 2012) under the
 238 Bayesian information criterion (BIC). The uncorrelated relaxed-clock mode and
 239 birth-death process were applied. The most recent common ancestors (TMRCA) of
 240 Ranunculales was constrained to a minimum age of 112 million years ago (Ma)
 241 according to the flower fossil assigned to *Teixeiraea lusitanica* von Balthazar, Pedersen
 242 & Friis (von Balthazar *et al.*, 2005; Magallón *et al.*, 2015). We set a minimum of 72 Ma
 243 to constrain the diverging between Berberidaceae and Ranunculaceae (Anderson *et al.*,
 244 2005; Sun *et al.*, 2018). Following the study of Bell *et al.*, (2010), we modeled both
 245 fossil calibrations as an exponential distribution (Ho & Phillips, 2009) with a mean of 1

246 and an offset (hard bound constraint) that equaled the minimum age of the calibrations.
 247 The MCMC was ran for 100 million generations sampling every 1,000 generations.
 248 Tracer v1.7.1 (Rambaut *et al.*, 2018) was used to assess the effective sample size (ESS >
 249 200) of each parameter. A maximum clade credibility tree was built by TreeAnnotator
 250 v2.5.2 (Rambaut & Drummond, 2010) using median node heights, with the initial 20%
 251 of trees discarded as burn-in.

252 Second, we estimated divergent time between haplotypes of *C. agrestis* using SNP
 253 sequences. All the parameters of BEAST2 were kept the same as above, with the
 254 exception of the TMRCA of *C. agrestis* constrained based on the previous results. We
 255 inferred ancestral distributions of *C. agrestis* using both statistical
 256 dispersal-extinction-cladogenesis (S-DEC) (Ree & Sanmartín, 2009) and statistical
 257 dispersal vicariance (S-DIVA) (Yu *et al.*, 2010) analyses. Both analyses were
 258 implemented in Reconstruct Ancestral State in Phylogenies (RASP) v4.0 (Yu *et al.*,
 259 2015), using 50,000 pruned trees (only *C. agrestis* included) from posterior distribution
 260 generated by BEAST2. For both analyses, the number of maximum areas at each node
 261 was set to two, and other parameters were left as default. Five biogeographic regions
 262 were defined according to the floristic division of China proposed by Wu & Wu (1998)
 263 and the phylogeographic study of Lin *et al.*, (2018): A, Qinling-Daba Mountains; B,
 264 North Hengduan Mountains; C, North QTP; D, South Hengduan Mountains; E, East
 265 Himalayan. To detect potential demographic expansions, temporal changes in the
 266 effective population size (N_e) were inferred with Extended Bayesian skyline plots
 267 (EBSP) as implemented in BEAST2 (Heled & Drummond, 2008). We used the SNP
 268 sequences of 139 individuals and the default substitution model. The MCMC was set to
 269 50 million generations sampling every 5,000 generations, and the first 20% was
 270 discarded as burn-in. Final graphs were produced with the custom R script provided by
 271 Heled (2010), with a burn-in cutoff of 20%.

272

273 **Effects of heterogeneous landscapes on genetic structure**

274 To illustrate the effects of the heterogeneous landscapes on shaping genetic structure,
 275 we conducted Mantel tests for the presence of IBD and IBR. To avoid the potential
 276 correlation between geographical distance and resistance distance, a partial Mantel test
 277 was also conducted to test IBD and IBR by controlling the resistance and geographical
 278 distance. GenAlEx v6.5 (Peakall & Smouse, 2012) was employed to compute the
 279 pairwise geographic distance among 18 populations. Resistance distance was generated
 280 in CIRCUITSCAPE v4.0.5 based on circuit theory (McRae, 2006; McRae *et al.*, 2008).
 281 Circuit resistance provide models of connectivity or resistance to predict patterns of
 282 dispersal among sites in a heterogeneous landscape (McRae, 2006; Dickson *et al.*,
 283 2019). The current ecological niche model (ENM) was calculated in MAXENT
 284 (Phillips & Dudik, 2008). We used 19 bioclimatic variables available from WorldClim2
 285 (Fick & Hijmans, 2017) at 30 arcseconds resolution (Table S1), and actual
 286 evapotranspiration (AET; <http://www.physicalgeography.net/fundamentals/8j.html>).
 287 To avoid multicollinearity, we ran a Pearson correlation analysis to eliminate one of the
 288 variables in each pair with a correlation value higher than 0.9. A total 12 climatic layers
 289 was retained for analyses (Table S1).

290 We used environmental niches representing connectivity between populations as
 291 conductance grids to produce pairwise resistance distance, as high habitat suitability
 292 was assumed to have low resistance (Nowakowski *et al.*, 2015). Tests of the
 293 significance for the relationship between geographical/resistance distances and genetic
 294 distance among populations were implemented in *ade4* v1.7
 295 (<https://CRAN.R-project.org/package=ade4>) using *mantel.rtest* with 9,999
 296 permutations.

297

298 **Environmental variables and genetic structure**

299 To estimate the contributions of environmental variables to genetic differentiation and
 300 to understand the turnover of allele frequencies along an environmental gradient, a
 301 gradient forest (GF) analysis was performed using *gradientForest* v0.1

(<http://gradientforest.r-forge.r-project.org/>). GF is a nonparametric, machine learning regression tree approach that allows for exploration of nonlinear associations of spatial, environmental, and allelic variables (Gugger *et al.*, 2017; Bay *et al.*, 2018). The analysis partitions the allele frequency data at split values along the environmental gradients and defines the amount of variation explained as ‘split importance’ values (Jiang *et al.*, 2019). The overall predictor importance plot of GF shows the mean importance weighted by allele R^2 , and the cumulative plot for each predictor shows cumulative change along the environmental gradient. The split importance values are added cumulatively to produce a step-like curve, thus, areas with large steps in a row indicate significant influence on allelic change. The 12 climatic layers used in above analysis was employed as environmental variables in the GF analysis.

A redundancy analysis (RDA) was performed to understand the associations between genetic structure and environmental variables. We estimated the proportion of genetic variance in the populations that is explained by environmental variables using six important variables identified by GF (bio03: isothermality; AET: actual evapotranspiration; bio04: temperature seasonality; bio07: temperature annual range; bio12: annual precipitation and bio15: seasonality precipitation, see results). We constrained the dependent variables (individuals) by the explanatory variables (climate). The RDA analysis was performed using the *rda* function in *vegan* v2.5 (Oksanen *et al.*, 2018; <http://CRAN.R-project.org/package=vegan>). The *anova.cca* function was used to calculate overall significance and significance of each climate variable was calculated using 9,999 permutations.

Genome-Environment Association (GEA)

To access the evidence of divergent selection acting on the genome and the genetic signatures of local adaptation, tests of associated outlier loci with environmental variation were carried out using BAYESCENV v1.1 (de Villemereuil & Gaggiotti, 2015), which extends the capabilities of the BAYESCAN algorithm by including a

330 model that incorporates environmental data. BAYESCENV has been shown to be
 331 fairly robust to isolation by distance and a hierarchically structured scenario (de
 332 Villemereuil & Gaggiotti, 2015). For a comprehensive consideration of the
 333 environmental effect, the environmental covariable layer was calculated with principal
 334 component analysis (PCA) of the six important bioclimatic variables, and the first PC
 335 axis (PC1), which explained most of the variability (48.62%), was extracted as the
 336 input for BAYESCENV. We ran 20 pilot runs of 5,000 iterations and a burn-in of
 337 50,000 iterations, ultimately obtaining 5,000 MCMC iterations for the analysis.
 338 Diagnostics of the log likelihoods and F_{ST} values for the 5,000 sampled iterations were
 339 checked using *coda* v0.19 (Plummer *et al.*, 2006) to confirm convergence and sample
 340 sizes of at least 2,500.

341 For functional annotation, the locus consensus sequences of all 6,120 SNPs were
 342 exported in *fasta* format. Candidate loci identified in GEA were annotated with
 343 Blast2Go v2 (Gotz *et al.*, 2008) by searching the non-redundant *Arabidopsis thaliana*
 344 protein database with records from the NCBI ESTs databases (Blastx). All the
 345 parameters were set as default with e-value < 1.0E-5.

346

347 **Results**

348 **Sequence data processing**

349 We produced 543 gigabyte (Gb) data containing 845,396,236 raw reads for 139
 350 individuals of *C. agrestis* (Table S2). After filtering, the average number of used reads
 351 per sample was 4,542,944 (median: 4,883,567; Table S2). The depths of coverage for
 352 processed samples ranged from 5.77x (WLG-2) to 22.44x (HZ2-3), with a mean
 353 coverage of 11.99x (Table S2). We optimized the parameters for STACKS analysis and
 354 obtained the optimal parameters for our dataset to be $M = n = 3$ (Fig. S2). The genome
 355 size of *C. agrestis* was estimated to be 1Gb using flow cytometry. After *de novo*
 356 assembling, we obtained 3,640,060 loci with mean genotyped sites per locus -144.49bp
 357 (a total of c.a. 0.51Gb), which comprises more than half of the genome. After strict

358 filtering ($-r = 1$), we obtained 6,959 RAD loci containing 6,120 variant sites that were
 359 used for population genetic analyses. The mean genotyped sites per locus was 146.59bp
 360 (stderr: 0.03). After a looser filtering ($-r = 0.8$), we obtained 24,247 RAD loci
 361 containing 22,915 variant sites. The sizes of 20 sequenced plastomes from 18
 362 populations of *C. agrestis* ranged from 150,979 bp (ZF) to 151,056 bp (SNJ) possessing
 363 the same 79 protein coding genes arranged in the same order (Table S3).

364

365 **Genetic diversity and structure**

366 We detected an extraordinary difference of genetic diversity at population level,
 367 especially for nucleotide diversity (π), which ranged from 0.0008 (HZ2) to 0.1646
 368 (DDL). The expected heterozygosity (H_e) and observed heterozygosity (H_o) were
 369 0.0007 (HZ2) - 0.1555 (DDL) and 0.0014 (HZ2) - 0.2650 (DDL; Table 1), respectively.
 370 When a loose filtering strategy was employed, we obtained a similar genetic diversity
 371 pattern (Table S4). In total, 66 haplotypes were detected (Fig. 1; Table S5) and the
 372 haplotype diversity (H_d) was 0.9798. Genetic differentiation between YL and HZ3 is
 373 the highest ($F_{ST}=0.9913$), and that between ZG and DDL is the lowest ($F_{ST}=0.0640$;
 374 Table S6; Fig. S3a). Estimates of N_e (Table S7) for each population ranged from 0.2
 375 (SNJ, TSM, XLS, DDL) to 1.2 (ZF). Six sites (HZ2, HZ3, KMG, YL, ZG and GS) had
 376 negative estimates and infinite 95% CI, which may result from either a truly large N_e or
 377 the consequence of limited sampling (Do *et al.*, 2014).

378 We conducted a Bayesian clustering analysis to explore the genetic structure of *C.*
 379 *agrestis*. The delta- K method identified the best-fit number (highest ΔK value) was two
 380 (Fig. S4a), referred to the East and West clades hereafter; the cross-entropy criterion
 381 indicated a better number of clusters was six (Fig. S4b). Considering the high F_{ST}
 382 among populations and the potential bias of ΔK method (Janes *et al.*, 2017), we further
 383 subdivided the two clades into two and four groups (E1, E2, W1, W2, W3 and W4),
 384 respectively, which was also supported by the PoCA (Fig. 2; Fig. S5). The AMOVA
 385 revealed that 77.59% of the overall variation was distributed among six groups, with

11.27% explained by variation among populations within groups (Table S8). Within genetic groups, W2 exhibited the highest nucleotide diversity (π ; 0.1646), while W3 had the lowest (0.0206). All genetic groups had lower observed heterozygosity than expected, with the exception of W2. In general, the East clade had lower levels of observed heterozygosity (H_o , E1=0.0029, E2=0.0034, W1=0.0548, W2=0.2650, W3=0.0134, W4=0.0088) and higher levels of inbreeding coefficient (F_{IS} , E1=0.1952, E2=0.1011, W1=0.0384, W2=-0.1806, W3=0.0227, W4=0.0510) than the West clade (Table 1). Both F_{ST} and D_{XY} showed similar patterns of divergence (Table S9 and S10; Fig. S3b), with genetic differentiation between W3 and W4 the highest (F_{ST} =0.8908; D_{XY} = 0.4307), followed by E2 and W4 (F_{ST} =0.8451; D_{XY} = 0.3670), and that between E1 and E2 the lowest (F_{ST} =0.1284; D_{XY} = 0.0620). Estimates of N_e ranged from 0.2 in W2 to 2.1 in E1 (Table S7). The G-statistic test suggested a significant population structure (p = 0.01**; Fig. S3c).

Population divergence

The 47-taxa 79-CDS matrix of Ranunculales used for estimating the original divergence of *C. agrestis* was 66,117 bp in length containing 12,998 parsimony-informative sites. The divergence of *C. agrestis* from *Kingdonia* was estimated to occur in the early Eocene (stem age, ca. 52.19 Ma; 95% highest posterior density [HPD] intervals, 26.12-83.13 Ma), whereas the divergence within *C. agrestis* started in the late Miocene (ca. 6.33 Ma; 95% HPD, 1.96-23.35 Ma) (Fig. S6). The tree topology inferred by BEAST2 using the SNP sequences identified two main clades (East and West clades) and six sub-groups (E1, E2, W1, W2, W3 and W4 groups), consistent with STRUCTURE and PoCA results. The divergence of the two main clades was estimated at appropriately 6.71 Ma (95% HPD, 5.18-7.89 Ma). The subsequent divergence between the six groups occurred mainly during the Pliocene to late Pleistocene (ca. 5.05Ma-2.90Ma) (Fig. 3a). Biogeographic analyses based on both S-DEC and S- DIVA supported Hengduan Mountains and adjacent regions as the most

likely ancestral region for *C. agrestis* (Fig. 3a, 3b and Fig. S7). For the demographic inference, EBSP analysis showed no sign of population expansion or contraction in demographic history of *C. agrestis* (Fig. 3c).

Effects of topography and ecology

Given that the distribution range of *C. agrestis* is topographically complex, we evaluated the influence of the heterogeneous landscape in shaping genetic structure via Mantel tests. Our analyses suggested both significant patterns of IBD ($R^2 = 0.443$, $P = 0.001^{**}$) and IBR ($R^2 = 0.315$, $P = 0.017^{*}$) (Fig. 4a and b), which were also supported by a partial Mantel test ($R_1^2 = 0.426$, $P = 0.001^{**}$; $R_2^2 = 0.192$, $P = 0.011^{*}$). Geographical distance explained the genetic structure better than resistance distances based on current distribution of habitats. We further plotted the conductance grid derived from the IBR model to illustrate the connectivity among populations and thus predict the probabilities of successful gene flow through a complex landscape. In general, connectivity gaps among populations were observed in the western clade, reflecting increased habitat isolation in this area resulted from complex topography. In contrast, locations of eastern clade had low resistance to dispersal thus high connectivity among populations (Fig. 4c).

Environmental variables associated with genetic structure

Among the 12 variables used for GF analysis, isothermality (bio03) was indicated as the most important predictor. Actual evapotranspiration (AET), temperature annual range (bio07) and temperature seasonality (bio04) were also of high importance. Seasonality precipitation (bio15), and annual precipitation (bio12) showed moderate importance to allele frequencies. The other six environmental variables had lower contributions (Fig. 5a). Fig. 5b shows the cumulative change in overall allele frequencies, where changes evolved with the environmental gradient.

The RDA analysis revealed a significant amount of genetic variation among

populations associated with the six important environmental variables (46.58%, $p = 0.001$). Each of the two axes explained a significant amount of variation (Axis 1: 58.55%, $P = 0.001^{**}$; Axis 2: 27.31%, $P = 0.001^{**}$; Fig. 6). All six environmental variables were ran separately and found to be significant. The contribution on genetic variation of each variable was generally consistent with the GF analysis. Isothermality (bio03; 18.93%, $P = 0.001^{**}$) was the most important predictor, whereas the temperature annual range (bio07) was less important in RDA analysis (Table 2).

Genome-Environment Association (GEA)

The BAYESCENV result revealed a total of 61 loci having a significant association with environmental variables. Sixteen of the 61 loci were successfully annotated by Blast2Go with cut-off of e-value $<1E-5$ and a mean similarity more than 80 (Table 3; Table S11). Among these were genes related to abiotic stress response and stress-induced morphological adaptations, such as *UGT74E2* (Loci ID: 2763; Gene Ontology (GO) term: transferase activity) regulating auxin homeostasis (Tognetti *et al.*, 2010), *SRK2E* (Loci ID: 8038; GO: protein kinase activity, ATP binding) involved in drought resistance (Mustilli *et al.*, 2002) and *AHK5* (Loci ID: 203056; GO: phosphorelay sensor kinase activity, phosphorelay signal transduction system) regulating stomatal state and transmitting the stress signal (Pham *et al.*, 2012), as well as genes required for the formation and development of leaves and flowers (*CYP71*, loci ID: 293398; GO: regulation of flower development, meristem structural organization leaf formation; *CSII*, loci ID: 218068; GO: cellulose biosynthetic, pollen tube development) (Li & Luan, 2011). These loci also include genes related to oxidation-reduction process involved in cellular respiration, for example *NDUFS7* (Loci ID: 83859; GO: oxidoreductase activity, NAD binding) and *NDUFB3* (Loci ID: 219291; GO: electron transport chain).

Discussion

Effects of topography on genetic structure

Our study presents the genetic structure of the palaeoendemic alpine *C. agrestis*. High haplotype diversity, uneven distribution of genetic diversity, as well as significant IBD and IBR patterns suggest restricted gene flow is likely a key factor in the high genetic differentiation among *C. agrestis* populations. The distribution area of *C. agrestis* is remarkable for its extraordinarily topographic diversity with some of the world's most rugged mountain ranges and altitudinal gradients spanning over 5,000 m (Wen *et al.*, 2014; Favre *et al.*, 2015). As a direct consequence of the uplift of the QTP, habitat isolation induced by complex topography and frequented climatic fluctuation was invoked as an important mechanism for the high diversity of alpine plants in this area (Wen *et al.*, 2014; Sun *et al.*, 2017). Hence, dramatic terrain and climate changes probably influenced genetic connectivity and shaped current genetic structure of *C. agrestis*.

We explored the number of genetic clusters of *C. agrestis* using delta-*K*, cross-entropy criterion and PCoA analysis. Despite potential bias of delta-*K* method which exhibits a pathology toward $K=2$ (Janes *et al.* 2017), the two distinct clades identified may reflect the early diversification of *C. agrestis*, and correspond to two unique climatic regions (Chen *et al.*, 2018). Both cross-entropy criterion and PCoA analysis indicated fine-scale structure within *C. agrestis*. For conservation purposes, fine-scale genetic structure can provide essential information for *in situ* conservation (Chung *et al.*, 2005). In particular, common garden experiments were found to be infeasible for *C. agrestis* during our field investigation, suggesting *ex situ* conservation strategies impractical.

Although previous studies suggested that drastically different levels of genetic diversity among populations may produce a bias estimate of F_{ST} (Charlesworth, 1998), both F_{ST} and D_{XY} showed similar pattern of genetic differentiation, revealing high isolation in the western clade. Coupled with the significant IBR pattern, habitat isolation, perhaps caused by complex topography, may be the cause for limited gene

flow as well as observed high differentiation. The QTP orogeny likely triggered the vicariance among different *C. agrestis* groups, consistent with the hypothesis of diverse isolated heterogeneous habitats on sky islands (He & Jiang, 2014), which has been reported in other alpine species (e.g. *Ligularia vellerea*, Asteraceae, Yang et al., 2012; *Salix brachista*, Salicaceae, Chen et al., 2019; *Ficus tikoua*, Moraceae, Deng et al., 2020). The north-south mountain chains and valleys may provide corridors for flora exchange between the north and south but represent barriers to migration between the east and west (Chen et al., 2018); likely resulting in the observed meridional differentiation of *C. agrestis* in the western clade as seen in previous studies (e.g., Gao et al., 2007; Li et al., 2011). Notably, the populations located in the Hengduan Mountains, exhibit high levels of genetic diversity (Table 1), which has been previously observed in plant species, such as *Parasynclathium souliei* (Asteraceae) (Lin et al., 2018), *Taxus wallichiana* (Taxaceae) (Liu et al., 2013) and *Quercus aquifolioides* (Fagaceae) (Du et al., 2017). As the QTP blocks the cold and dry air (Liu et al., 2013), the Hengduan Mountains are characterized by a much warmer and wetter climate, providing a comparatively stable environment. Although the western populations are likely to possess higher genetic diversity, they are also likely to be more vulnerable to shifting disturbance regimes and environmental changes due to lack of genetic connectivity resulted from habitats isolation. Compared to the western clade, north-south mountain barriers among different latitudes may be driving the divergence among the groups in the eastern clade. A conductance grid derived from the IBR model showed low resistance to dispersal, thus high connectivity, among populations in the eastern clade. In addition, E1 and E2 exhibit a relatively low level of genetic differentiation. This phenomenon is likely due to the similar local environments among different populations resulting from the quite moderate topographical and climatic characteristics in this region (Chen et al., 2018).

The concept of effective size is key to conservation genetics, as it summarizes population status regarding inbreeding and genetic drift and provides the prospects for

the sustainability of the population (Wang *et al.*, 2016). Although N_e is frequently less than the census population size, small values of N_e combined with low heterozygosity suggest high levels of inbreeding in populations of *C. agrestis*. The reduction of genetic diversity, as well as the fixation of mildly deleterious alleles that reduce reproductive fitness, influences long-term population persistence and poses a risk of extinction (Kramer & Havens, 2009; Oakley & Winn, 2012; Poudel *et al.*, 2014). In addition, episodes of low population size have a disproportionate effect on the overall value of N_e (Charlesworth, 2009). Thus, low N_e may also be attributed to the disturbance caused by human activities or possible bottleneck events in the evolutionary history of *C. agrestis*.

Impact of environmental heterogeneity in population divergence

Our molecular dating based on plastome sequences suggested an early Eocene origin of *C. agrestis*, which is accordant with previous estimates (Ruiz-Sanchez *et al.*, 2012) and recent results from nuclear data (Sun *et al.*, 2020). The diversification of *C. agrestis* has increased since the Pliocene (Fig. 3), which is likely related to intense uplift of the Hengduan Mountains during the late Miocene to late Pliocene (Mulch & Chamberlain, 2006; Sun *et al.*, 2011; Favre *et al.*, 2015). The environmental and climatic fluctuations triggered by mountain uplifts can always lead to divergent selection and species adaptation associated with dramatic ecological niche changes (Ren *et al.*, 2017). Orogenic activities not only provide numerous new habitats favorable for diversification, but also promote population divergence due to hindered gene flow. Our biogeographic analysis revealed that the demographic evolution of *C. agrestis* has not simply been influenced by the orogeny at a particular time in history, but involves repeated episodes of dispersal and subsequent vicariance events. Coupled with the evidence of significant IBD and IBR patterns, we speculate natural selection driven by extraordinary environmental heterogeneity and geomorphological dynamics may have profound impact in the diversification of *C. agrestis*. In addition, during our field work we found the census population sizes of *C. agrestis* in the Himalaya-Hengduan

Mountains to be much smaller than other sites. Previous studies have indicated a significant negative effect of small population size on adaptive evolution (Nevado *et al.*, 2019). Hence, we speculate that natural selection may play double-edged roles in the evolutionary history of any given species. On one hand, heterogeneous environments could provide more new ecological niches conducive for speciation, on the other hand, it may put the populations with weak adaptive ability to new environment in danger of extinction. Furthermore, significant IBR pattern suggest that genetic structure would be better explained by combining habitat isolation between populations than just geographical distance alone. Thus, habitat changes may have large effect on genetic structure for *C. agrestis*, indicating its sensitivity to environmental niche.

Studying local adaptation contributes to understanding the ability of populations to sustain or adapt to rapid climate change (Jia *et al.*, 2019). We associated environmental variables with genetic structure to demonstrate the impact of environmental heterogeneity on population divergence. Specifically, we found a significant association between genetic variation and temperature variables (Fig. 4; bio03, AET, bio04, bio07), suggesting temperature is an important driver of genetic variation within *C. agrestis*, as detected in previous studies of angiosperm trees (Ahrens *et al.*, 2019; Jia *et al.*, 2019; Jiang *et al.*, 2019). In fact, there is growing evidence that high-mountain environments experience more rapid changes in temperature than environments at lower elevations (Pepin *et al.*, 2015; Palazzi *et al.*, 2019). Despite that, little attention has been given to understanding the mechanisms of adaptive response to climate change in high altitude areas. *C. agrestis* is habitually confined to grow under trees, shrubs or rocks, indicating sensitivity to temperature. Moreover, the vertical distribution of *C. agrestis* is extensive, ranging from 2,100 to 5,000 meters above sea level (Fu & Bartholomew, 2001), exhibiting a considerable differentiation in temperature. Therefore, it is reasonable that temperature variables play a vital role in promoting genetic divergence within *C. agrestis* and given the ongoing climate warming, temperature is likely to be a key driver for adaptation of *C.*

582 *agrestis* in the future.

583

584 **Genomic signatures associated with local adaptation**

585 BAYESCV identified a set of genes that may be involved in a continuously
586 adaptive process. Some genes such as pivotal kinases involved in signaling pathway
587 regulation, may induce abiotic stress responses for defense in harsh climates in the
588 QTP. For example, genes associated with abscisic acid (ABA) signaling pathway
589 (*SRK2EA*) can regulate numerous ABA responses, such as stomata closure, to
590 response to high altitudes with extreme difference in temperature (Sierla *et al.*, 2018).
591 Considering the tiny leaves of *C. agrestis*, we speculate that stomatal regulation would
592 be a vital process for maintaining hydration and regulating temperature. As altitude
593 increases, hypobaric hypoxia becomes the main factor interfering with life activity.
594 Genes related to oxidation-reduction processes involved in cellular respiration, such
595 as *NDUFS7* and *NDUFB3*, are of importance in facilitating adaptation of *C. agrestis* to
596 high altitude areas of hypoxia. In addition, some putatively adaptive genes (*CYP71*
597 and *CSII*) are associated with physiological trait regulation of vegetative and
598 reproductive organs (Gu *et al.*, 2010). Thus, morphological adaptations may be an
599 essential alternative for *C. agrestis* to respond to environmental stress.

600 A recent study characterizing the genome of *Kingdonia uniflora* (sister to *C.*
601 *agrestis*) calculated the nucleotide substitution rate of 1.4×10^{-9} per site per year (Sun
602 *et al.*, 2020), which is comparatively lower than an estimate of *Arabidopsis thaliana*
603 (7.0×10^{-9}) (Ossowski *et al.*, 2010) and indirect estimates based on the divergence
604 between monocots and dicots ($5.8-8.1 \times 10^{-9}$) (Wolfe *et al.*, 1987). The low nucleotide
605 substitution rate may be evidence of low speciation rate, which likely contributes to
606 the paucity of species in the family and the low morphological diversity of *C. agrestis*
607 and *K. uniflora*. Low evolution rate is also indicative of low mutation rate, reducing
608 the standing genetic variation which is the genetic source of adaptation (Lai *et al.*,
609 2019). In isolated populations with high homozygosity and little standing genetic

variation, only large effect alleles can escape the effects of drift becoming the targets of selection (Orr, 1998; Rausher & Delph, 2015; Sella & Barton, 2019). Additionally, many methods for detecting adaptive genetic variation often only have the power to detect loci and alleles with large phenotypic effects (Wellenreuther & Hansson, 2016; Luikart et al., 2018). Nonetheless, genetic patterns that confer adaptation to environmental heterogeneity are mostly polygenic and controlled by numerous genetic variants of small effect (Savolainen *et al.*, 2013; Ahrens *et al.*, 2019). Therefore, the genes identified as putatively adaptive in *C. agrestis* should constitute the first step in studying genetic mechanism of local adaptation in alpine environments. Further work characterizing the whole genome is needed to obtain more precise functions of genes under selective pressure and illuminate patterns of potential polygenic adaptation (Mayol *et al.*, 2019). Generally, these patterns can serve as a proxy for understanding and monitoring the adaptive process in rapidly changing environment (Ahrens *et al.*, 2019), while informing future conservation management.

Conclusions

In this study, we demonstrate two major advances. First, we provide new insights into the genetic structure and evolutionary history of a relictual alpine herb surviving in heterogeneous environment, which can inform further conservation. Our results shed light on vital roles of mountain uplift in promoting population diversification and the dual effect of environmental heterogeneity on species evolutionary history. Second, our analyses provide a roadmap to study evolution of local adaptation of non-model species for which common garden experiments are infeasible and genomic resources are limited. By conducting multiple analyses, we associated environmental variables with genetic variation and explicitly measured the relative influence of bioclimate on population divergence. In turn, the important environmental predictors were implemented in identifying potential loci that may under divergent selection to gain in-depth knowledge of genetic basis involved in local adaptation. More importantly,

our results improve our understanding of drivers and mechanisms of adaptive evolution, aiding efforts on developing sound conservation programs in facing of rapid changing climate.

Acknowledgments

We are grateful to Jie Cai for help with sample collection in Tibet. We thank Jialiang Li and Kangshan Mao for help with the calculation of absolute differentiation (D_{XY}). This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA20050203), the Programme Foundation for the Backbone of Scientific Research by Wuhan Botanical Garden, Chinese Academy of Sciences (Y855241G01), the Major Program of National Natural Science Foundation of China (31590823), and the National Key R and D Program of China (2017YFC0505200).

Author contributions

HCW, HS, and YXS developed the idea and designed the experiment. XZ, JWZ, LSY, NL, HJZ, RG, LJL, YHZ and TD collected the leaf materials. XZ and YXS performed the statistical analyses; XZ, YXS, JBL, HS and HCW interpreted the results and wrote the manuscript. All authors read, edited and approved the final manuscript. XZ and YXS contributed equally to this work.

Data accessibility

All newly sequenced plastomes were deposited in National Center for Biotechnology Information (NCBI) with accession numbers MT228704-MT228722. RAD raw read data are stored at the NCBI Sequence Read Archive in the Bioproject PRJNA616150. Genomic data sets and R scripts used in this study are available in the DRYAD archives under accession doi: 10.5061/dryad.4f4qrfj7p.

Reference

666 **Ahrens CW, Byrne M, Rymer PD. 2019.** Standing genomic variation within coding and
667 regulatory regions contributes to the adaptive capacity to climate in a foundation tree
668 species. *Mol Ecol* **28**(10): 2502-2516.

669 **Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. 2008.** Adaptation,
670 migration or extirpation: climate change outcomes for tree populations. *Evol Appl* **1**(1):
671 95-111.

672 **Anderson CL, Bremer K, Friis EM. 2005.** Dating phylogenetically basal eudicots using *rbcL*
673 sequences and multiple fossil reference points. *American Journal of Botany* **92**:
674 1737-1748.

675 **Bay RA, Harrigan RJ, Underwood VL, Gibbs HL, Smith TB, Ruegg K. 2018.** Genomic
676 signals of selection predict climate-driven population declines in a migratory bird.
677 *Science* **359**(6371): 83.

678 **Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina
679 sequence data. *Bioinformatics* **30**(15): 2114-2120.

680 **Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A,
681 Drummond AJ. 2014.** BEAST 2: a software platform for Bayesian evolutionary
682 analysis. *PLoS Comput Biol* **10**(4): e1003537.

683 **Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013.** Stacks: an analysis
684 tool set for population genomics. *Molecular ecology* **22**(11): 3124-3140.

685 **Charlesworth B. 1998.** Measures of divergence between populations and the effect of forces
686 that reduce variability. *Molecular Biology and Evolution* **15**: 538-543.

687 **Charlesworth B. 2009.** Effective population size and patterns of molecular evolution and
688 variation. *Nature Reviews Genetics* **10**(3): 195-205.

689 **Chen J, Huang Y, Brachi B, Yun Q, Zhang W, Lu W, Li H, Li W, Sun X, Wang G. et al.
690 2019.** Genome-wide analysis of Cushion willow provides insights into alpine plant
691 divergence in a biodiversity hotspot. *Nat Commun* **10**: 5230.

692 **Chen Y-S, Deng T, Zhou Z, Sun H. 2018.** Is the East Asian flora ancient or not? *National
693 Science Review* **5**(6): 920-932.

694 **Chung MY, Suh Y, López-Pujol J, Nason JD, Chung MG. 2005.** Clonal and fine-scale
695 genetic structure in populations of a restricted Korean endemic, *Hosta jonesii*
696 (Liliaceae) and the implications for conservation. *Ann Bot.* **96**:279-288.

697 **Colautti RI, Barrett SCH. 2013.** Rapid Adaptation to Climate Facilitates Range Expansion of
698 an Invasive Plant. *Science* **342**(6156): 364.

699 **Conover DO, Duffy TA, Hice LA. 2009.** The Covariance between Genetic and Environmental
700 Influences across Ecological Gradients. *Annals of the New York Academy of Sciences*
701 **1168**(1): 100-129.

702 **Coop G, Witonsky D, Di Rienzo A, Pritchard JK. 2010.** Using Environmental Correlations
703 to Identify Loci Underlying Local Adaptation. *Genetics* **185**(4): 1411.

704 **Coyne JA, Orr HA. 2004.** *Speciation*. Sunderland, MA, USA: Sinauer Associates.

705 **Darriba D, Taboada GL, Doallo R, Posada D. 2012.** jModelTest 2: more models, new
706 heuristics and parallel computing. *Nat Methods* **9**(8): 772.

707 **Darwin C. 1859.** *On the origin of species by means of natural selection, or the preservation of*

- 708 *favoured races in the struggle for life*. London, UK: John Murray.
- 709 **de Villemereuil P, Gaggiotti OE. 2015.** A new FST-based method to uncover local adaptation
710 using environmental variables. *Methods in Ecology and Evolution* **6**(11): 1248-1258.
- 711 **Deng JY, Fu R-H, Compton SG, Liu M, Wang Q, Yuan C, Zhang L-S, Chen Y. 2020.** Sky
712 islands as foci for divergence of fig trees and their pollinators in southwest China. *Mol*
713 *Ecol.* **29**:762–782.
- 714 **Dickson BG, Albano CM, Anantharaman R, Beier P, Fargione J, Graves TA, Gray ME,**
715 **Hall KR, Lawler JJ, Leonard PB, et al. 2019.** Circuit-theory applications to
716 connectivity science and conservation. *Conservation Biology* **33**(2): 239-249.
- 717 **Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. 2014.** NeEstimator V2:
718 re-implementation of software for the estimation of contemporary effective population
719 size (*N_e*) from genetic data. *Molecular Ecology Resources* **14**(1): 209-214.
- 720 **Du F, Hou M, Wang W, Mao K, Hampe, A. 2017.** Phylogeography of *Quercus aquifolioides*
721 provides novel insights into the Neogene history of a major global hotspot of plant
722 diversity in south-west China. *J. Biogeogr.* **44**: 294-307.
- 723 **Earl DA, vonHoldt BM. 2012.** STRUCTURE HARVESTER: a website and program for
724 visualizing STRUCTURE output and implementing the Evanno method. *Conservation*
725 *Genetics Resources* **4**(2): 359-361.
- 726 **Ellegren H. 2014.** Genome sequencing and population genomics in non-model organisms.
727 *Trends in Ecology & Evolution* **29**(1): 51-63.
- 728 **Excoffier L, Laval G, Schneider S. 2007.** Arlequin (version 3.0): an integrated software
729 package for population genetics data analysis. *Evolutionary bioinformatics online* **1**:
730 47-50.
- 731 **Favre A, Packert M, Pauls SU, Jahnig SC, Uhl D, Michalak I, Muellner-Riehl AN. 2015.**
732 The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas.
733 *Biol Rev Camb Philos Soc* **90**(1): 236-253.
- 734 **Fick SE, Hijmans RJ. 2017.** WorldClim 2: new 1-km spatial resolution climate surfaces for
735 global land areas. *International Journal of Climatology* **37**(12): 4302-4315.
- 736 **Forester BR, Lasky JR, Wagner HH, Urban DL. 2018.** Comparing methods for detecting
737 multilocus adaptation with multivariate genotype–environment associations.
738 *Molecular Ecology* **27**(9): 2215-2233.
- 739 **Foster AS. 1971.** Additional studies on the morphology of blind vein-endings in the leaf of
740 *circaeaster agrestis*. *American Journal of Botany* **58**(3): 263-272.
- 741 **Frichot E & Francois O. 2015.** LEA: an R package for Landscape and Ecological Association
742 studies. *Methods in Ecology and Evolution* **6**:925-929.
- 743 **Frichot E, Mathieu F, Trouillon T, Bouchard G, Francois O. 2014.** Fast and efficient
744 estimation of individual ancestry coefficients. *Genetics* **196**:973-983.
- 745 **Friis G, Fandos G, Zellmer AJ, McCormack JE, Faircloth BC, Milá B. 2018.**
746 Genome-wide signals of drift and local adaptation during rapid lineage divergence in a
747 songbird. *Molecular Ecology* **27**(24): 5137-5153.
- 748 **Fu D, Bartholomew B 2001.** CIRCAEASTERACEAE. In: Wu ZY, Raven, P. H. & Hong, D. Y.
749 ed. *Flora of China*. Beijing & St. Louis: Science Press & Missouri Botanical Garden

Press, 439.

Gao LM, MÖller M, Zhang XM, Hollingsworth ML, Liu J, Mill RR, Gibby M, Li DZ. 2007. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Molecular Ecology* **16**(22): 4684-4698.

Goicoechea PG, Guillardín L, Fernández-Ibarrodo L, Valbuena-Carabaña M, González-Martínez SC, Alía R, Kremer A. 2019. Adaptive Introgression Promotes Fast Adaptation In Oaks Marginal Populations. *bioRxiv*: 731919.

Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* **36**(10): 3420-3435.

Goudet J, Raymond M, de Meeüs T, Rousset F. 1996. Testing Differentiation in Diploid Populations. *Genetics* **144**(4): 1933.

Gu Y, Kaplinsky N, Bringmann M, Cobb A, Carroll A, Sampathkumar A, Baskin TI, Persson S, Somerville CR. 2010. Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. *Proceedings of the National Academy of Sciences* **107**(29): 12866.

Gugger PF, Liang CT, Sork VL, Hodgskiss P, Wright JW. 2017. Applying landscape genomic tools to forest management and restoration of Hawaiian koa (*Acacia koa*) in a changing environment. *Evolutionary Applications* **11**(2): 231-242.

He K, Jiang XL. 2014. Sky islands of southwest China. I: An overview of phylogeographic patterns. *Chinese Science Bulletin* **59**: 585-597.

Heled J. 2010. Extended Bayesian Skyline Plots Tutorial.

Heled J, Drummond AJ. 2008. Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* **8**(1): 289.

Hill WG. 1981. Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* **38**:209-216.

Ho SYW, Phillips MJ. 2009. Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. *Systematic Biology* **58**(3): 367-380.

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. 2016. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist* **188**(4): 379-397.

Hughes AC. 2017. Understanding the drivers of Southeast Asian biodiversity loss. *Ecosphere*, **8**(1): e01624.

Hughes CE & Atchison GW. 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.* **207**: 275-282.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**(14): 1801-1806.

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI. 2017. The K

791 = 2 conundrum. *Mol Ecol.* **26**: 3594-3602.

792 **Jia K-H, Zhao W, Maier PA, Hu X-G, Jin Y, Zhou S-S, Jiao S-Q, El-Kassaby YA, Wang T,**
793 **Wang X-R, et al. 2019.** Landscape genomics predicts climate change-related genetic
794 offset for the widespread *Platycladus orientalis* (Cupressaceae). *Evolutionary*
795 *Applications* <https://doi.org/10.1111/eva.12891>.

796 **Jiang X-L, Gardner EM, Meng H-H, Deng M, Xu G-B. 2019.** Land bridges in the
797 Pleistocene contributed to flora assembly on the continental islands of South China:
798 Insights from the evolutionary history of *Quercus championii*. *Molecular*
799 *Phylogenetics and Evolution* **132**: 36-45.

800 **Jombart T. 2008.** adegenet: a R package for the multivariate analysis of genetic markers.
801 *Bioinformatics* **24**(11): 1403-1405.

802 **Katoh K, Standley DM. 2013.** MAFFT multiple sequence alignment software version 7:
803 improvements in performance and usability. *Mol Biol Evol* **30**(4): 772-780.

804 **Kawecki TJ, Ebert D. 2004.** Conceptual issues in local adaptation. *Ecology Letters* **7**(12):
805 1225-1241.

806 **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper**
807 **A, Markowitz S, Duran C, et al. 2012.** Geneious Basic: an integrated and extendable
808 desktop software platform for the organization and analysis of sequence data.
809 *Bioinformatics* **28**(12): 1647-1649.

810 **Kimura M, Crow JF. 1964.** The number of alleles that can be maintained in a finite population.
811 *Genetics* **49**(4): 725-738.

812 **Kramer AT, Havens K. 2009.** Plant conservation genetics in a changing world. *Trends in Plant*
813 *Science* **14**(11): 599-607.

814 **Lai Y, Yeung KL, Omland E, Pang EL, Liao BY, Cao HF, Zhang BW, Yeh CF, Hung CM,**
815 **Hung HY, et al. 2019.** Standing genetic variation as the predominant source for
816 adaptation of a songbird. *Proc Natl Acad Sci USA* **116**: 2152-2157.

817 **Li H, Luan S. 2011.** The cyclophilin AtCYP71 interacts with CAF-1 and LHP1 and functions
818 in multiple chromatin remodeling processes. *Mol Plant* **4**(4): 748-758.

819 **Li J, Milne RI, Ru D, Miao J, Tao W, Zhang L, Xu J, Liu J, Mao K. 2020.** Allopatric
820 divergence and hybridization within *Cupressus chengiana* (Cupressaceae), a
821 threatened conifer in the northern Hengduan Mountains of western China. *Molecular*
822 *Ecology* doi:10.1111/mec.15407.

823 **Li Y, Zhai S-N, Qiu Y-X, Guo Y-P, Ge X-J, Comes HP. 2011.** Glacial survival east and west
824 of the ‘Mekong–Salween Divide’ in the Himalaya–Hengduan Mountains region as
825 revealed by AFLPs and cpDNA sequence variation in *Sinopodophyllum hexandrum*
826 (Berberidaceae). *Molecular Phylogenetics and Evolution* **59**(2): 412-424.

827 **Lin N, Deng T, Moore MJ, Sun Y, Huang X, Sun W, Luo D, Wang H, Zhang J, Sun H.**
828 **2018.** Phylogeography of *Parasyncalathium souliei* (Asteraceae) and Its Potential
829 Application in Delimiting Phylogeoregions in the Qinghai-Tibet Plateau
830 (QTP)-Hengduan Mountains (HDM) Hotspot. *Front. Genet.* **9**:171.

831 **Liu J, Möller M, Provan J, Gao L-M, Poudel RC, Li D-Z. 2013.** Geological and ecological
832 factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytologist*

199(4): 1093-1108.

Lowry DB, Lovell JT, Zhang L, Bonnette J, Fay PA, Mitchell RB, Lloyd-Reilley J, Boe AR, Wu Y, Rouquette FM, et al. 2019. QTL $\times$ environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. *Proceedings of the National Academy of Sciences* **116**(26): 12933.

Lu HY & Guo ZT. 2014. Evolution of the monsoon and dry climate in East Asia during late Cenozoic: a review. *Science China Earth Sciences* **57**: 70–79.

Luikart G, Kardos M, Hand BK, Rajora OP, Aitken SN, Hohenlohe PA. 2018. Population genomics: advancing understanding of nature. In: Rajora OP, ed. *Population genomics: concepts, approaches and applications*. Cham, Switzerland: Springer, 3–79.

Mayol M, Riba M, Cavers S, Grivet D, Vincenot L, Cattonaro F, Vendramin G, González-Martínez SC. 2019. A multiscale approach to detect selection in nonmodel tree species: Widespread adaptation despite population decline in *Taxus baccata* L. *Evol Appl.* **00**:1–18.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* **66**(2): 526-538.

McRae BH. 2006. Isolation by resistance. *Evolution* **60**:1551–1561.

McRae BH, Dickson BG, Keitt TH, Shah VB. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* **89**:2712–2724.

Mulch A & Chamberlain CP. 2006. Earth science – The rise and growth of Tibet. *Nature* **439**: 670–671.

Mustilli AC, Merlot S, Vavasseur A, Fenzi F, Giraudat J. 2002. Arabidopsis OST1 protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reactive oxygen species production. *The Plant cell* **14**(12): 3089-3099.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858.

Nevado B, Wong ELY, Osborne OG, Filatov DA. 2019. Adaptive Evolution Is Common in Rapid Evolutionary Radiations. *Current Biology* **29**: 3081–3086.

Noroozi J, Talebi A, Doostmohammadi M, Rumpf SB, Linder HP, Schneeweiss GM. 2018. Hotspots within a global biodiversity hotspot – Areas of endemism are associated with high mountain ranges. *Scientific Reports* **8**: 10345

Nowakowski AJ, DeWoody JA, Fagan ME, Willoughby JR, Donnelly MA. 2015. Mechanistic insights into landscape genetic structure of two tropical amphibians using field-derived resistance surfaces. *Mol Ecol* **24**(3): 580-595.

Oakley CG & Winn AA. 2012. Effects of population size and isolation on heterosis, mean fitness, and inbreeding depression in a perennial plant. *New Phytologist* **196**: 261-270.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, et al. 2018. Vegan: Community Ecology Package. <https://CRAN.R-project.org/package=vegan>.

Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* **52**:935-949.

875 **Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, et al. 2010.** The
876 rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*.
877 *Science* **327**:92-94

878 **Peakall R, Smouse PE. 2012.** GenAlEx 6.5: genetic analysis in Excel. Population genetic
879 software for teaching and research--an update. *Bioinformatics* **28**(19): 2537-2539.

880 **Pham J, Liu J, Bennett MH, Mansfield JW, Desikan R. 2012.** Arabidopsis histidine kinase 5
881 regulates salt sensitivity and resistance against bacterial and fungal infection. *New*
882 *Phytol* **194**(1): 168-180.

883 **Phillips SJ, Dudík M. 2008.** Modeling of species distributions with Maxent: new extensions
884 and a comprehensive evaluation. *Ecography* **31**(2): 161-175.

885 **Plummer M, Best N, Cowles K, Vines K. 2006.** CODA: convergence diagnosis and output
886 analysis for MCMC. *R News* **6**(1): 7-11.

887 **Poudel RC, Möller M, Liu J, Gao L-M, Baral SR, Li D-Z. 2014.** Low genetic diversity and
888 high inbreeding of the endangered yews in Central Himalaya: implications for
889 conservation of their highly fragmented populations. *Diversity and Distributions*
890 **20**(11): 1270-1284.

891 **Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using
892 multilocus genotype data. *Genetics* **155**(2): 945-959.

893 **Qu XJ, Moore MJ, Li DZ, Yi TS. 2019.** PGA: a software package for rapid, accurate, and
894 flexible batch annotation of plastomes. *Plant Methods* **15**(1): 50.

895 **R Team C 2014.** R: A language and environment for statistical computing. R Foundation for
896 Statistical Computing. Vienna, Austria.

897 **Rambaut A, Drummond A. 2010.** TreeAnnotator version 1.6. 1. *University of Edinburgh,*
898 *Edinburgh, UK.* Available at: <http://beast.bio.ed.ac.uk>.

899 **Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018.** Posterior Summarization
900 in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* **67**(5): 901-904.

901 **Rausher MD & Delph LF. 2015.** Commentary: When does understanding phenotypic
902 evolution require identification of the underlying genes? *Evolution* **69**(7): 1655–1664.

903 **Ree RH, Sanmartín I. 2009.** Prospects and challenges for parametric models in historical
904 biogeographical inference. *Journal of Biogeography* **36**: 1211-1220.

905 **Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R. 2015.** A practical guide to
906 environmental association analysis in landscape genomics. *Molecular Ecology* **24**(17):
907 4348-4370.

908 **Ren G, Mateo RG, Liu J, Suchan T, Alvarez N, Guisan A, Conti E, Salamin N. 2017.**
909 Genetic consequences of Quaternary climatic oscillations in the Himalayas: *Primula*
910 *tibetica* as a case study based on restriction site-associated DNA sequencing. *New*
911 *Phytologist* **213**(3): 1500-1512.

912 **Ren Y, Li Z, Hu Z. 2003.** Approaches to the systematic position of *Circaeaster* based on the
913 morphological data. *Acta Botanica Boreali-occidentalia Sinica* **23**(7): 1091-1097.

914 **Rochette NC, Catchen JM. 2017.** Deriving genotypes from RAD-seq short-read data using
915 Stacks. *Nature Protocols* **12**: 2640.

916 **Rosenberg NA. 2004.** Distruct: a program for the graphical display of population structure.

917 *Molecular Ecology Notes* **4**(1): 137-138.

918 **Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P,**

919 **Ramos-Onsins SE, Sanchez-Gracia A. 2017.** DnaSP 6: DNA Sequence

920 Polymorphism Analysis of Large Data Sets. *Mol Biol Evol* **34**(12): 3299-3302.

921 **Ru D, Sun Y, Wang D, Chen Y, Wang T, Hu Q, Abbott RJ, Liu J. 2018.** Population genomic

922 analysis reveals that homoploid hybrid speciation can be a lengthy process. *Molecular*

923 *Ecology* **27**(23): 4875-4887.

924 **Ruiz-Sanchez E, Rodriguez-Gomez F, Sosa V. 2012.** Refugia and geographic barriers of

925 populations of the desert poppy, *Hunnemannia fumariifolia* (Papaveraceae).

926 *Organisms Diversity & Evolution* **12**(2): 133-143.

927 **Savolainen O, Lascoux M, Merilä J. 2013.** Ecological genomics of local adaptation. *Nature*

928 *Reviews Genetics* **14**: 807.

929 **Savolainen O, Pyhäjärvi T, Knürr T. 2007.** Gene Flow and Local Adaptation in Trees.

930 *Annual Review of Ecology, Evolution, and Systematics* **38**(1): 595-619.

931 **Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, Peichel**

932 **CL, Saetre G-P, Bank C, Brännström Å, et al. 2014.** Genomics and the origin of

933 species. *Nature Reviews Genetics* **15**: 176.

934 **Sella G & Barton NH. 2019.** Thinking About the Evolution of Complex Traits in the Era of

935 Genome-Wide Association Studies. *Annu. Rev. Genom. Hum. Genet.* **20**:461-93

936 **Sierla M, Horak H, Overmyer K, Waszczak C, Yarmolinsky D, Maierhofer T, Vainonen**

937 **JP, Salojärvi J, Denessiouk K, Laanemets K, et al. 2018.** The Receptor-like

938 Pseudokinase GHR1 Is Required for Stomatal Closure. *The Plant cell* **30**(11):

939 2813-2837.

940 **Song JH, Kang HS, Byun YH, Hong SY. 2010.** Effects of the Tibetan Plateau on the Asian

941 summer monsoon: a numerical case study using a regional climate model.

942 *International Journal of Climatology* **30**: 743–759.

943 **Spieth PT. 1974.** Gene flow and genetic differentiation. *Genetics* **78**(3): 961.

944 **Stroud DA, Surgenor EE, Formosa LE, Reljic B, Frazier AE, Dibley MG, Osellame LD,**

945 **Stait T, Beilharz TH, Thorburn DR, et al. 2016.** Accessory subunits are integral for

946 assembly and function of human mitochondrial complex I. *Nature* **538**(7623): 123-126.

947 **Sun B-N, Wu J-Y, Liu Y-S, Ding S-T, Li X-C, Xie S-P, Yan D-F, Lin Z-C. 2011.**

948 Reconstructing Neogene vegetation and climates to infer tectonic uplift in western

949 Yunnan, China. *Palaeogeography, Palaeoclimatology, Palaeoecology* **304**(3):

950 328-336.

951 **Sun H, Zhang J, Deng T, Boufford DE. 2017.** Origins and evolution of plant diversity in the

952 Hengduan Mountains, China. *Plant Diversity* **39**(4): 161-166.

953 **Sun Y, Moore MJ, Landis JB, Lin N, Chen L, Deng T, Zhang J W., Meng AP, Zhang SJ,**

954 **Tojibaev OS, et al. (2018).** Plastome phylogenomics of the early-diverging eudicot

955 family Berberidaceae. *Mol. Phylogenet. Evol.* **128**: 203–211.

956 **Sun Y, Deng T, Zhang A, Moore MJ, Landis JB, Lin N, Zhang H, Zhang X, Huang J,**

957 **Zhang X, et al. 2020.** The draft genome of the endangered, relictual plant *Kingdonia*

958 *uniflora* (Circaeasteraceae, Ranunculales) reveals potential mechanisms and perils of

- 959 evolutionary specialization. *bioRxiv*: 2020.2001.2008.898460.
- 960 **Sun Y, Moore MJ, Lin N, Adelalu KF, Meng A, Jian S, Yang L, Li J, Wang H. 2017.**
- 961 Complete plastome sequencing of both living species of Circaeasteraceae
- 962 (Ranunculales) reveals unusual rearrangements and the loss of the *ndh* gene family.
- 963 *BMC Genomics* **18**(1): 592.
- 964 **Tang H, Micheels A, Eronen JT, Ahrens B & Fortelius M. 2013.** Asynchronous responses of
- 965 East Asian and Indian summer monsoons to mountain uplift shown by regional climate
- 966 modelling experiments. *Climate Dynamics* **40**: 1531–1549.
- 967 **Tognetti VB, Van Aken O, Morreel K, Vandenbroucke K, van de Cotte B, De Clercq I, et**
- 968 **al. 2010.** Perturbation of indole-3-butyric acid homeostasis by the
- 969 UDP-glucosyltransferase UGT74E2 modulates Arabidopsis architecture and water
- 970 stress tolerance. *The Plant cell*, **22**(8), 2660-2679.
- 971 **von Balthazar M, Pedersen K, Friis E. 2005.** *Teixeiraea lusitanica*, a new fossil flower from
- 972 the Early Cretaceous of Portugal with affinities to the Ranunculales. *Plant Systematics*
- 973 *and Evolution* **255**: 55–75.
- 974 **Wang J, Santiago E, Caballero A. 2016.** Prediction and estimation of effective population
- 975 size. *Heredity* **117**: 193-206.
- 976 **Waples RS. 2006.** A bias correction for estimates of effective population size based on linkage
- 977 disequilibrium at unlinked gene loci. *Conservation Genetics* **7**:167-184.
- 978 **Weigel D, Nordborg M. 2015.** Population Genomics for Understanding Adaptation in Wild
- 979 Plant Species. *Annual Review of Genetics* **49**(1): 315-338.
- 980 **Weir BS, Cockerham CC. 1984.** Estimating F-statistics for the analysis of population
- 981 structure. *Evolution* **38**(6): 1358-1370.
- 982 **Wellenreuther M, Hansson B. 2016.** Detecting polygenic evolution: problems, pitfalls, and
- 983 promises. *Trends Genet.* **32**:155–64.
- 984 **Wen J, Zhang JQ, Nie ZL, Zhong Y, Sun H. 2014.** Evolutionary diversifications of plants on
- 985 the Qinghai-Tibetan Plateau. *Front Genet* **5**: 4.
- 986 **Wolfe KH, Li WH, Sharp PM. 1987.** Rates of nucleotide substitution vary greatly among
- 987 plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* **84**:
- 988 9054-9058.
- 989 **Wright S. 1946.** Isolation by distance under diverse systems of mating. *Genetics* **31**(1): 39-59.
- 990 **Wu ZY, Wu SG. 1996.** A proposal for a new floristic kingdom (realm): the E. Asiatic Kingdom,
- 991 its delineation and characteristics. In: Zhang AL, Wu SG, eds. *Floristic characteristics*
- 992 *and diversity of East Asian plants*. Beijing, China: China Higher Education
- 993 Press/Springer, 3–42.
- 994 **Yang JB, Li DZ, Li HT. 2014.** Highly effective sequencing whole chloroplast genomes of
- 995 angiosperms by nine novel universal primer pairs. *Mol Ecol Resour* **14**(5): 1024-1031.
- 996 **Yang ZY, Yi TS, Pan YZ Gong X. 2012.** Phylogeography of an alpine plant *Ligularia vellerea*
- 997 (Asteraceae) in the Hengduan Mountains. *J Syst Evol* **50**:316–324
- 998 **Yu Y, Harris AJ, Blair C, He X. 2015.** RASP (Reconstruct Ancestral State in Phylogenies): a
- 999 tool for historical biogeography. *Mol Phylogenet Evol* **87**: 46-49.
- 1000 **Zhang D, Gao F, Jakovlic I, Zou H, Zhang J, Li WX, Wang GT. 2019.** PhyloSuite: An

1001 integrated and scalable desktop platform for streamlined molecular sequence data
1002 management and evolutionary phylogenetics studies. *Mol Ecol Resour*
1003 doi:10.1111/1755-0998.13096.

Table 1 Summary of statistics calculated for the 6120 variant positions. n, number of genotype; π , average nucleotide diversity; H_e , average expected heterozygosity per locus; H_o , average observed heterozygosity per locus; the inbreeding coefficients (F_{IS}). The statistics of six defined groups are bolded. E1, eastern group1; E2, eastern group2; W1, western group1; W2, western group2; W3, western group3; W4, western group4

Group IDs	Population IDs	n	π	H_e	H_o	F_{IS}
E1	-	75	0.0464	0.0461	0.0029	0.1952
	SNJ	13	0.0207	0.0199	0.0038	0.0383
	TBS	17	0.0218	0.0212	0.0029	0.0532
	HZ2	9	0.0008	0.0007	0.0014	-0.0010
	HZ3	10	0.0013	0.0012	0.0022	-0.0017
	KMG	8	0.0015	0.0014	0.0027	-0.0023
	TSM	9	0.0214	0.0203	0.0027	0.0425
	WLG	4	0.0306	0.0287	0.0034	0.0621
	XLS	5	0.0223	0.0210	0.0026	0.0375
E2	-	19	0.0384	0.0374	0.0034	0.1011
	BLS	9	0.0216	0.0204	0.0035	0.0529
	SGN	10	0.0254	0.0241	0.0033	0.0573
W1	-	7	0.0688	0.0639	0.0548	0.0384
	MXG	4	0.0215	0.0189	0.0220	-0.0005
	YL	1	0.0016	0.0008	0.0016	-
	ZG	2	0.1633	0.1225	0.1470	0.0244
W2	DDL	9	0.1646	0.1555	0.2650	-0.1806
W3	SJL	10	0.0206	0.0196	0.0134	0.0227
W4	-	19	0.0249	0.0242	0.0088	0.0510
	YH	8	0.0150	0.0140	0.0057	0.0238
	GS	1	0.0255	0.0128	0.0255	-
	ZF	10	0.0136	0.0130	0.0095	0.0100

Table 2 RDA results based on six important environmental variables identified by GF analysis.

Environmental variables	Constrained proportion	F	p
Bio03	0.1893	31.98	0.001**
AET	0.1712	28.31	0.001**
Bio04	0.1408	22.45	0.001**
Bio12	0.0902	13.59	0.001**
Bio07	0.0627	9.17	0.001**
Bio15	0.0203	2.84	0.001**

1011 **Table 3** Annotation information of 16 candidate loci under selection that the E-value of blastx is
1012 less than 1E-5. Detailed functional annotation is provided in Table S11.

Loci ID	Gene names	Description	Length	E-Value	Similarity
2763	<i>UGT74E2</i>	UDP-glycosyltransferase 74E2-like	146	1.16E-19	82.08
8038	<i>SRK2E</i>	Serine/threonine-protein kinase TIO-like isoform X1	146	4.97E-07	81.56
8150	<i>CSCI</i>	Calcium permeable stress-gated cation channel 1-like	146	5.12E-20	92.71
19591	<i>Atlg18270</i>	Ketose-bisphosphate aldolase class-II family protein	146	6.34E-06	87.07
83859	<i>NDUFS7</i>	NADH dehydrogenase subunit 7 (mitochondrion)	146	7.55E-26	93.96
126986	<i>RTL</i>	Retrotransposon gag protein	147	4.78E-14	84.5
146932	<i>RPPL1</i>	Putative disease resistance RPP13-like protein 1 isoform X1	146	2.78E-06	87.52
203056	<i>AHK5</i>	Histidine kinase 5	146	7.55E-14	84.61
206146	<i>ABCC2</i>	ABC transporter C family member 2-like	146	2.77E-15	89.16
218068	<i>CSII</i>	Protein cellulose synthase interactive 1	146	3.74E-14	86.88
219291	<i>NDUFB3</i>	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 3-B-like	146	8.26E-20	82.13
243120	<i>NUDT8</i>	Nudix hydrolase 8	146	8.49E-14	81.72
275249	<i>N/A</i>	Glycosyl transferase (glycosyl transferase family 2)	146	3.38E-24	85.59
293398	<i>CYP71</i>	Peptidyl-prolyl cis-trans isomerase CYP71	146	7.81E-17	92.11
296907	<i>N/A</i>	Glucan endo-1,3-beta-glucosidase 13-like	146	1.45E-20	83.33
305334	<i>CSLC6</i>	Probable xyloglucan glycosyltransferase 6	146	7.31E-26	98.54

1013

1014

Figure legends

Fig. 1 Geographical distribution of haplotypes detected in polymorphic sites of *Circaeaster agrestis* populations. Colors indicate genetic groups based on STRUCTURE and PoCA results (see also Fig. 2), with the Eastern clade being made up of the orange and dark green circles, with the Western clade consisting of the light green, light pink, purple, and blue circles. Circle sizes reflect the number of individuals sampled.

Fig. 2 Population genetic structure in *Circaeaster agrestis* based on 6,120 SNPs from (a) PCoA for all populations, east-clade and west-clade, corresponding percentage of variation explained by each principal component is provided in **Fig. S5**; and (b) STRUCTURE with K=2 (showing the highest ΔK ; Fig. S4a) and 6 (representing the divergence within east and west clades).

Fig. 3 Evolutionary history of *Circaeaster agrestis*. (a) Divergent time estimation and ancestral area reconstructions of 66 haplotypes using BEAST and statistical dispersal-extinction-cladogenesis (S-DEC), respectively. Mean divergence dates and 95% HPDs for major nodes (1–5) are summarized in top left. Pie charts on each node indicate marginal probabilities for each alternative ancestral area derived from S-DEC. Results are based on a maximum area number of two. Inferred dispersal and vicariance events are indicated by green and red circle respectively. (b) Five biogeographic regions representing the current distributions of *C. agrestis*, according to the floristic division of China proposed by Wu & Wu (1998) and the study of Lin et al., (2018): A, Qinling-Daba Mountains; B, North Hengduan Mountains; C, North QTP; D, South Hengduan Mountains; E, East Himalayan regions. (c) Temporal changes in the effective population size (N_e) changes inferred with Extended Bayesian skyline plots (EBSP) using BEAST2. Re, reproductive number, δ , death rate.

Fig. 4 Topographic and ecological effects on genetic structure. Relationship of genetic distance and (a) geographical distance, (b) resistance distance based on climatic niche suitability, as tested by a Mantel test. (c) Conductance grid derived from the Isolation-

1043 by-Resistance (IBR) model. Grid with warm colour indicates high conductance, cold
1044 colour indicates high resistance.
1045 **Fig. 5** (a) R²-weighted importance of environmental variables that explain genetic
1046 gradients from gradient forest (GF) analysis. (b) Cumulative importance of allelic
1047 change along the first six environmental gradients. The units of environmental variables
1048 are provided in the parentheses. SD, standard deviation.

1049 **Fig. 6** Redundancy analysis showing the relationship between the independent climate
1050 parameters and population structure. Individuals are colored points and colors represent
1051 six groups (E1, E2, W1, W2, W3, W4). Small black points are SNPs. The definition of
1052 each climate variable is provided in the top right corner.

1053

Supporting information

1055 **Fig. S1** The known distribution area of the *Circaeaster agrestis* based on all recorded
1056 sampling points available in Chinese Virtual Herbarium (CVH; <http://www.cvh.ac.cn/>).
1057 The green triangles indicate sampling locations in this study. The red cross indicates a
1058 distribution record in reported in 1997, but not available now based on one recent field
1059 investigation in 2018.

1060 **Fig. S2** (a) The number of 80% polymorphic loci shared across most samples (the r80
1061 loci) as M=n increases. (b) The distribution of the number of SNPs per locus for a range
1062 of M=n values.

1063 **Fig. S3** (a) F_{ST} values within populations estimated by *hierfstat*, Group names of
1064 populations were labeled in the top of boxes. (b) F_{ST} values within genetic groups
1065 estimated by *hierfstat*. (c) The results of G-statistic test.

1066 **Fig. S4** The optimal K value identified using (a) delta- K method in STRUCTURE
1067 HARVESTER, and (b) cross-entropy criterion in R package *LEA*.

1068 **Fig. S5** Scree plot of the percentage of variation explained by each principal component
1069 (PC), corresponding to **Fig. 2a** for all populations, East-clade and West-clade,
1070 respectively.

1071 **Fig. S6** Divergence times of Ranunculales estimated by BEAST2 with a relaxed
1072 molecular clock based on the combined protein-coding region sequences. A and B
1073 indicate fossil calibration points. C and D indicate the origination and diversification of
1074 *Circaeaster agrestis*, respectively. Median ages of nodes are shown with bars indicating
1075 the 95% highest posterior density intervals for each node.

1076 **Fig. S7** Ancestral area reconstructions of 66 haplotypes using statistical dispersal
1077 vicariance (S-DIVA) analysis. Pie charts on each node indicate marginal probabilities
1078 for each alternative ancestral area derived from S-DIVA. Results are based on a
1079 maximum area number of two. Inferred dispersal and vicariance events are indicated by
1080 green and red circle respectively. Defined biogeographic regions are the same as the
1081 statistical dispersal-extinction-cladogenesis analysis (Fig. 3).

1082 **Table S1** Summary of population information of *Circaeaster agrestis*, including the
1083 number of individuals genotyped at each localities (n), geographic information, voucher
1084 specimens information, and 20 bioclimate values (bio01-bio19 and AET) of each
1085 location. AET, the actual evapotranspiration. The units of environmental variables are
1086 provided in the parentheses. SD, standard deviation.

1087 **Table S2** Numbers of reads of RAD-seq for each individual of *Circaeaster agrestis*.

1088 **Table S3** Assemble and annotation information of newly sequenced plastomes of
1089 *Circaeaster agrestis* populations.

1090 **Table S4** Summary of statistics calculated for the 22915 variant positions. n, number of
1091 genotype; π , average nucleotide diversity; H_e , average expected heterozygosity per
1092 locus; H_o , average observed heterozygosity per locus; the inbreeding coefficients (FIS).
1093 The statistics of six defined groups are bolded. E1, eastern group1; E2, eastern group2;
1094 W1, western group1; W2, western group2; W3, western group3; W4, western group4.

1095 **Table S5** Haplotypes information detected in 18 populations based on 6120 SNPs.

1096 **Table S6** Paired population-level F_{ST} estimated by *hierfstat* using Weir and
1097 Cockerham's method.

1098 **Table S7** Estimated N_e values using linkage disequilibrium method, implemented in

1099 NeEstimator, with a minor allele frequency cutoff of 0.05 and 95% confidence intervals
1100 (CI) estimated by jackknifing. The statistics of six defined groups are bolded. E1,
1101 eastern group1; E2, eastern group2; W1, western group1; W2, western group2; W3,
1102 western group3; W4, western group4; n, number of genotype.

1103 **Table S8** The analysis of molecular variance (AMOVA) for SNP data among six genetic
1104 groups (E1, E2, W1, W2, W3 and W4)

1105 **Table S9** Paired group-level F_{ST} estimated by *hierfstat* using Weir and Cockerham's
1106 method.

1107 **Table S10** Paired mean absolute differentiation (D_{XY}) among genetic groups using a
1108 Perl script provided by Ru et al., (2018).

1109 **Table S11** Detailed annotation information of sixteen genes under potential divergent
1110 selection. GO: Gene Ontology

(a)

All populations

East-clade

West-clade

(b)

East-clade

West-clade

$K = 2$ **$K = 6$**

E1

E2

W1

W2

W4

W3

(a)

(b)

(c)

(a)

(b)