Landscape connectivity of *Cercidiphyllum japonicum*, an endangered species and its implications for conservation

Bi-cai Guan [a,*], Wei Chen [a], Xi Gong [a], Ting Wu [a], Qi-ying Cai [a], Yi-zhen Liu [a], Gang Ge [b,a,*]

[a] *College of Life Sciences, Nachang University, Nachang 330031, China*
[b] *Key Laboratory of Poyang Lake Environment and Resource Utilization, Nanchang University, Ministry of Education, Nanchang 330031, China*

ABSTRACT

Cercidiphyllum japonicum, a Tertiary relict, recolonized areas north of the Yangtze River after the last glacial; however, little is known about its specific colonization corridors. Together with distribution models, the least cost path (LCP) analysis has been used to reveal the landscape connectivity of species. In this study, we utilized the categorical LCP method, combining the species distribution with genetic data from cpDNA and nuclear markers, to identify the possible dispersal routes of *C. japonicum* after the LGM. Across time periods and genetic markers, the results revealed that the species generally spread from the western edge of the Sichuan Basin, while the highest degree of dispersal potential corresponds with the year 2080 and the cpDNA haplotype. Furthermore, shifts in the species' range and the indication of an area of low genetic divergence further support the existence of a dispersal corridor. Overall, we believe that a dispersal route from the western edge of the Sichuan Basin through the Qinling Mountains and further to the northeast could exist, and therefore, the results are an important supplement to the evolutionary history of *C. japonicum*. In the future, we believe species distribution models (SDM) and connectivity assessment in relation to climate change will provide increasingly useful information and new implications for prioritizing the conservation of the endangered species.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The use of ecological information and historical climatic and environmental data to guide the construction of appropriate phylogenetic and demographic models has added to our understanding of the role of particular geological barriers and climatic changes in intraspecific divergence (Chan et al., 2011). Recent approaches have been building upon the use of species distribution models (SDMs) by coupling them with coalescent analyses based on genetic data that allow for a better understanding of how demographic processes impact (or impacted) populations, both now and in the past (Knowles and Alvarado-Serrano, 2010; Brown and Knowles, 2012; Gehara et al., 2013; He et al., 2013; Dellicour et al., 2014). With the advent of the use of habitat heterogeneity as a friction layer with geospatial and environmental data, several approaches such as the least cost path (LCP) method and circuit theory developed using GIS-based methods or data have played an important role in many fields of biogeography (Vignieri, 2005; McRae and Beier, 2007; Graves et al., 2014). Notably, the dispersal network tool can be used to visualize dispersal corridors across landscapes and identify biogeographic barriers (Chan et al., 2011; Zeller et al., 2012; Brown, 2014). The GIS-based method can be useful and effective in revealing important information relating

to evolution and conservation, especially in regard to endangered species (Brown and Yoder, 2015; Huerta-Ramos et al., 2015).

The cool-temperate deciduous forests in mainland China were generally believed to have retreated southward at the LGM. Since the Holocene, they expanded their range northward due to warmer and wetter conditions into the previously uninhabitable northern region, reaching their maximum geographical range around the mid-Holocene (Fang, 1991; Wang and Sun, 1994; Yu et al., 2000; Harrison et al., 2001; Ni et al., 2006). Recently, growing phylogeography studies have localized recurring glacial refugia in this region, such as at the Hengduan Mountains, the Qinling Mountains, and the Daba Mountains (Zhang, Z.Y., et al., 2006; Zhang, X.L., et al., 2006; Li et al., 2011; Qiu et al., 2011). However, little is still known about the recolonization route of the cool-temperate deciduous forests during the late Quaternary. The main objective of this paper is to localize their dispersal corridors during the late Quaternary using a model organism.

Cercidiphyllum japonicum is a tall canopy tree whose range extends widely eastward from southwest of the Sichuan Basin to the coastline of eastern China at a latitude between 25°N and 36°N; it is usually found in small discrete populations (Isagi et al., 2005; Krassilov, 2010). The species is well-known for its shortage of seedlings in its habitat resulting from relatively low seed fertility. In recent years, due to over-exploitation and human disturbance, the species' populations have diminished sharply and become further fragmented, and it is now on the edge of extinction in China (Wan and Zhang, 2002). The

* Corresponding authors.
E-mail addresses: guanbicai12@163.com (B. Guan), gge@ncu.educn (G. Ge).

preliminary SDMs data in conjunction with molecular evidence from previous work on the species strongly suggest its recolonization of areas north of the Yangtze after the last glacial — that is, following the retreat of the arid steppe and desert vegetation there (Winkler and Wang, 1993; Yu et al., 2000; Harrison et al., 2001). This is the principal reason why we select *C. japonicum* as a model to identify their dispersal corridors. Moreover, previous work on the maternal chloroplast (cp) and biparental nuclear (n) DNA of the species emphasized various inheritance and dispersal features. Using landscape genetics approaches, we expected to be able to infer more comprehensive and convincing dispersal corridors. In addition, fossil evidence identified *Cercidiphyllum* as a dominant component of Tertiary forest and woody pioneer communities throughout the Northern Hemisphere (Mai, 1995; Meyer and Manchester, 1997; Manchester et al., 2009). Substantial fossil records would provide actual evidence for the dispersal hypothesis.

Here, we aimed to estimate the distributional changes of *C. japonicum* and identify its dispersal corridors during the late Quaternary. Furthermore, we expected to reveal some valuable implications for the evolutionary history and conservation priority of the endangered plant.

2. Methods

2.1. Determination of the shared haplotypes

When the integrated LCP method is used to localize the movement pathway based on molecular data, one premise is that populations with shared haplotypes usually experienced dispersal between sample localities (Chan et al., 2011; Yu et al., 2015). To explore the possible dispersal routes of *C. japonicum*, the populations that have shared haplotypes needed to be determined first. In a previous phylogeographical study, Qi et al. (2012) collected samples from 33 populations from mainland China and found 12 and 25 haplotypes for cpDNA and nDNA, respectively. Among the haplotypes, there were five shared haplotypes from each genome. For this study, we adopted the haplotype data from the 33 localities, and populations with shared haplotypes for the two types of molecular markers are shown in the haplotype distribution maps (Fig. 1).

2.2. SDM

We used climate and species occurrence data to build the species distribution models, identifying areas where *C. japonicum* could have existed or could exist in relation to three general time periods with their corresponding environmental conditions. The three general time periods are 1) the last glacial maximum (LGM, ~21–18 ka) based on the Community Climate System Model (CCSM4)), 2) the present (~1950–2000), and 3) the year 2080 based on CCSM4 with representative concentration pathway 85. Environmental data used in this study, included data of 19 bioclimate variables (Table S1), which were obtained from the WorldClim database (http://www.worldclim.org; Hijmans et al., 2005) and had a resolution of 30 arc-sec (ca. 1 km^2) for each layer.

To reduce over-fitting caused by redundant variables, we selected the most explanatory and uncorrelated variables, eliminating the others based on the highest and most significant correlation coefficients (Spearman correlation index, r > 0.75 and p < 0.05) (Schrag et al., 2008) (Table S2). We used the maximum entropy algorithm in MAXENT v3.3 (Phillips et al., 2006) to generate the three period distribution models of *C. japonicum.* Model performance was evaluated using the area under the (receiver operating characteristic) curve (AUC) calculated by MAXENT. Values between 0.75 and 0.9 indicate good discrimination (Swets, 1988). We modeled the distribution 10 times with cross-validation type, using a different 80% of localities to train the model and 20% to test the model, and visually compared AUC scores and jackknife tests of variable importance to assess consistency between runs. We used the default parameters and the following user-selected features: regularization multiplier of 2.0, application of a random seed, duplicate presence records removal and logistic probabilities used for the output (Phillips and Dudik, 2008). In addition, we performed a pairwise comparison of the binary SDMs for the three time periods (e.g., from the LGM to the present) to predict the range shifts since the LGM period into the future.

For the data necessary for the SDMs, the species occurrence localities were mainly gathered from the literature, the database of the Global Biodiversity Information Facility (GBIF, http://data.gbif.org), and the database of the Chinese Virtual Herbarium (CVH, http://cvh.org.cn). For the SDMs to be performed well, occurrence data were required to be spatially independent, so the elimination of spatially clustered localities was important for model calibration and evaluation. When clusters of localities exist, models often over-fit predictions because of environmental biases, which reduces the model's predictive ability in relation to the spatially independent data because the model performance values are inflated (Veloz, 2009; Hijmans, 2012; Boria et al., 2014). Hence, a method that spatially rarefies the occurrence data at several distances was applied according to the climatic heterogeneity (Boria et al., 2014). The occurrence localities were spatially filtered, respectively, at 5, 15, and 25 km^2 in areas of high, medium and low climatic heterogeneity. For the estimation of climatic heterogeneity, we conducted a principal component (PC) analysis for the bioclimate variables, and then calculated the mean standard deviation of the first three climate
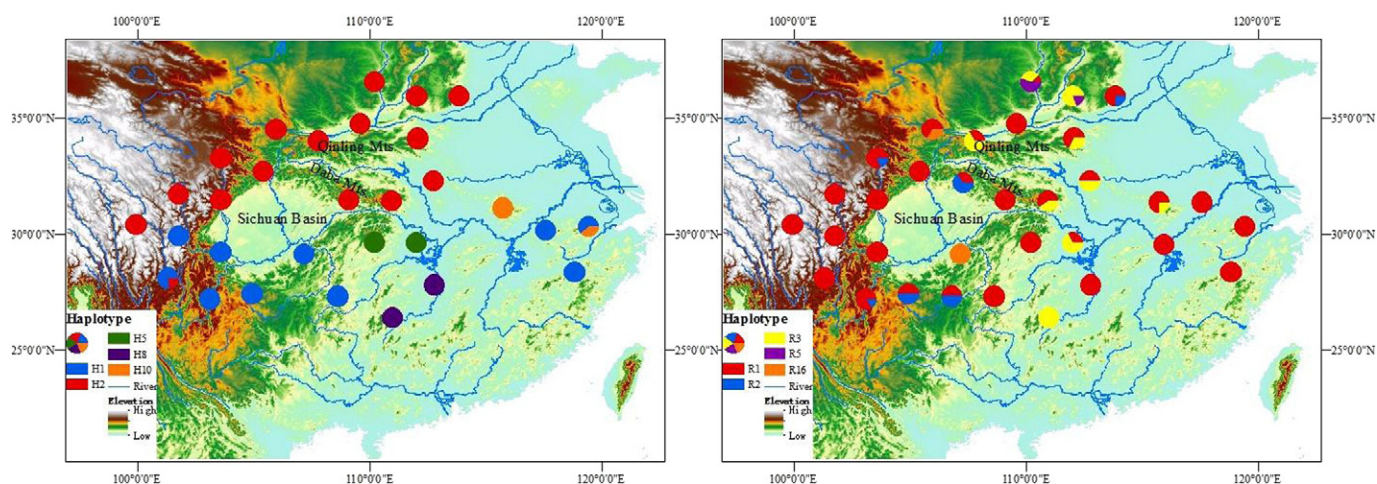

Fig. 1. Shared haplotypes distribution based on two genetic markers: (A) chloroplast (cp) DNA; (B) nuclear (n) DNA.