Molecular Ecology (2006) **15**, 4109–4122 doi: 10.1111/j.1365-294X.2006.03086.x

Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA

HONG-WEI WANG* and SONG GE*,†
*State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China,
†Graduate School, Chinese Academy of Sciences, Beijing 100039, China

Abstract

***Cathaya argyrophylla* is an endangered conifer restricted to subtropical mountains of China. To study phylogeographical pattern and demographic history of *C. argyrophylla*, species-wide genetic variation was investigated using sequences of maternally inherited mtDNA and biparentally inherited nuclear DNA. Of 15 populations sampled from all four distinct regions, only three mitotypes were detected at two loci, without single region having a mixed composition (G_{ST} = 1). Average nucleotide diversity (θ_{ws} = 0.0024; π_s = 0.0029) across eight nuclear loci is significantly lower than those found for other conifers (θ_{ws} = 0.003~0.015; π_s = 0.002~0.012) based on estimates of multiple loci. Because of its highest diversity among the eight nuclear loci and evolving neutrally, one locus (2009) was further used for phylogeographical studies and eight haplotypes resulting from 12 polymorphic sites were obtained from 98 individuals. All the four distinct regions had at least four haplotypes, with the Dalou region (DL) having the highest diversity and the Bamian region (BM) the lowest, paralleling the result of the eight nuclear loci. An AMOVA revealed significant proportion of diversity attributable to differences among regions (13.4%) and among populations within regions (8.9%). F_{ST} analysis also indicated significantly high differentiation among populations (F_{ST} = 0.22) and between regions (F_{ST} = 0.12–0.38). Non-overlapping distribution of mitotypes and high genetic differentiation among the distinct geographical groups suggest the existence of at least four separate glacial refugia. Based on network and mismatch distribution analyses, we do not find evidence of long distance dispersal and population expansion in *C. argyrophylla*. *Ex situ* conservation and artificial crossing are recommended for the management of this endangered species.**

Keywords: *Cathaya argyrophylla*, glacial refugia, mtDNA haplotypes, nuclear DNA sequences, phylogeography, population genetics

Received 25 March 2006; revision received 2 June 2006; accepted 23 June 2006

Introduction

Global climatic fluctuations, particularly Quaternary climate oscillations resulted in repeated drastic environmental changes and have profoundly shaped the distribution ranges and genetic structure of many animals and plants from different latitudes (Taberlet *et al*. 1998; Abbott *et al*. 2000; Avise 2000; Hewitt 2000). Although fossil pollen records and palaeoenvironmental data as well as the present distribution of genetic variability have been used for investigating the demographic history of the species (Comes & Kadereit 1998), molecular biogeography or phylogeography has provided additional sources of information on the glacial history of species and their range changes or shifts as inferred from the geographical patterns of genetic variation (Comes & Kadereit 1998; Avise 2000; Hewitt 2000, 2004; Schonswetter *et al*. 2005). To date, major effects on the genetic structure of environmentally induced range changes have been increasingly appreciated and the possible consequences have been outlined, in particular for species in Europe and North America (Soltis *et al*. 1997; Taberlet *et al*. 1998; Avise 2000;

Correspondence: Song Ge, Fax: +86-10-62590843, E-mail: gesong@ibcas.ac.cn

Hewitt 2000, 2004; Schonswetter *et al*. 2005). Compared with North America and Europe, however, relatively few studies have been conducted in plant and animal species in temperate montane and tropic areas (Hewitt 2000, 2004).

China has the most diverse flora of any country in the North Temperate zone, with the number of species of vascular plants being two times that of North America and three times that of Europe although China is nearly the same size as Europe or continental United States (Axelrod *et al*. 1996). One of the important reasons for the floristic richness of China is its tectonically active, highly dissected, elevated geography associated with the collision of the Indian subcontinent with the mainland of Asia, which has made the region not only an important centre of survival but also an important centre of speciation and evolution (Axelrod *et al*. 1996; Ying 2001). Although its most part has never been covered by ice sheets, China, together with its neighbouring areas in eastern Asia, has experienced the development of cooler and drier climates within the last 15-million year period (Axelrod *et al*. 1996). The tremendous climatic changes during this period, particularly Quaternary glaciations, have lead to most extinction and influenced the distribution and evolution of many plants in China and its neighbouring areas (Axelrod *et al*. 1996; Lu *et al*. 2001; Ge *et al*. 2002; Cheng *et al*. 2005; Shen *et al*. 2005; Zhang *et al*. 2005).

Cathaya argyrophylla Chun et Kuang (Pinaceae) has been categorized as a palaeoendemic, with a fossil history dating at least to the Cretaceous (Liu & Basinger 2000). As one of the most endangered conifers with total number of mature individuals less than 2000 (Sun *et al*. 1994), *C. argyrophylla* has been named as the 'giant panda' of the plant kingdom and occurs currently in widely separated subtropical areas in China (Wang 1990; Xie & Chen 1999; Liu & Basinger 2000). Because of its unique systematic position in the family Pinaceae and importance in the studies of palaeoclimate and palaeogeology (Fu 1992), this species has attracted many investigations on different aspects of its biology (Hu & Wang 1984; Wang 1990; Fu 1992; Sun *et al*. 1994; Xie *et al*. 1999; Liu & Basinger 2000). However, relatively few studies have been conducted on its population genetics and evolutionary history. Based on 25 allozyme loci, Ge *et al*. (1998) investigated the genetic diversity and population genetic structure of *C. argyrophylla* and found lower genetic variation, particularly at the population level, and significant higher population differentiations among regions and among populations within regions relative to other conifers (Ge *et al*. 1998). A similar genetic pattern has also been revealed by random amplified polymorphic DNA (RAPD) markers (Wang *et al*. 1997). Ge *et al*. (1998) speculated that historical factors such as severe bottleneck and subsequent genetic drift during Quaternary glaciations and limited gene flow in postglacial times were the main factors responsible for the unique population genetic structure of this species. However, our understanding of the population genetic structure at different geographical scales and of the evolutionary history such as potential refugia and population expansion in *C. argyrophylla* is still limited because previous markers such as allozyme and RAPDs are of limited utility in inferring population history and dynamics.

In recent decades, molecular markers have been used extensively to infer the biogeographical history of different organisms and to test phylogeographical hypotheses related to the presence of isolated glacial populations and refugia (see reviews in Schaal *et al*. 1998; Taberlet *et al*. 1998; Avise 2000; Hewitt 2004). The use of molecular markers derived from different genomes provides a more complete description of population structure and insights into population history and dynamics, particularly for comparisons of biparentally inherited nuclear and maternally inherited organelle markers (Schaal *et al*. 1998; Burban & Petit 2003; Petit *et al*. 2005). Here we report the results of a phylogeographical study on *C. argyrophylla* based on sequence data of the maternally inherited mitochondrial DNA (mtDNA) and biparentally inherited nuclear loci. We were particularly interested in the following questions: (i) Should the genetic profile revealed by the sequence data of mtDNA and nuclear fragments be consistent with that of previous allozyme and RAPD studies? (ii) What evolutionary factors might explain the patterns and levels of genetic variation observed? (iii) Do the current isolated populations represent the distinct glacial populations or refugia and is there evidence for recent population declines or expansions? Information on population genetics and demographical history of *C. argyrophylla* is also of practical importance because this endangered species has become a major conservation concern and there is an urgent need to develop recommendations for the management and recovery of this endangered species.

Materials and methods

Population sampling

Cathaya argyrophylla was formally described in 1958 (Chun & Kuang 1958) as the only extant member of the genus that was apparently distributed in North America and East Asia during the Cretaceous (Liu & Basinger 2000). Like most conifers, *C. argyrophylla* is a monoecious, wind-pollinated, and predominantly outcrossing species. It is a member of the needle-leaved and broad-leaved mixed forest in middle-elevation mountains of the subtropical China, where rainfall is abundant and the climate is cool in the summer and cold in the winter (Wang 1990; Fu 1992; Ge *et al*. 1998). At present, this species is confined to four widely separated mountainous regions, i.e. the Dalou mountains (DL), Bamian mountains (BM), Yuecheng

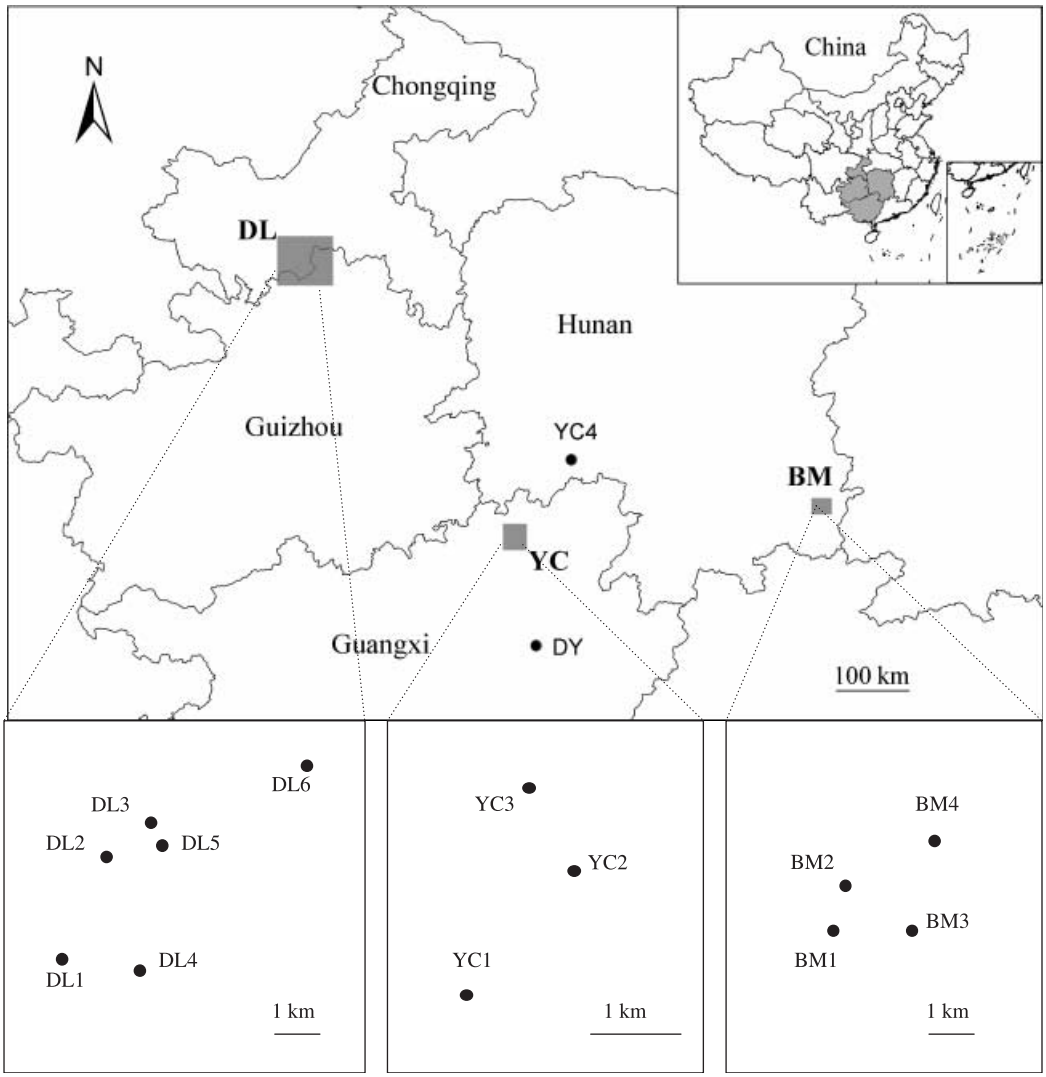

Fig. 1 Distribution map of *Cathaya argyrophylla* and the locations of populations sampled in this study. Populations were assigned to four geographical regions (BM, DL, DY, and YC) as described in the text.

mountains (YC), and Dayao mountains (DY) (Fig. 1), with population sizes ranging from one to several dozens (Xie 1995; Xie *et al*. 1999). In addition to its limited number of individuals and small population sizes, the species is characterized by unusually low fertility and very low survival and growth rates (Wang 1990; Xie & Chen 1999). On average, for example, there are only 4.3 seeds per cone, with as many as 12.2% of cones producing no seed; in addition, cone production is very low in nature and the percent of seed germination in the field is only 21% (Xie & Chen 1999). In extreme cases, not a single tree produces viable seeds in some populations (personal observations). Furthermore, much evidence suggests that the existing populations of *C. argyrophylla* in various communities are declining and are at risk of being replaced by fastgrowing, broad-leaved trees (Fu 1992; Xie & Chen 1999; Xie *et al*. 1999). Accordingly, *C. argyrophylla* is currently listed as one of the most endangered plant species in China (Fu 1992; Wang & Xie 2004).

During two field surveys from 1992 to 1994 and from 2002 to 2004, we have made extensive collections of cones throughout the range of *C. argyrophylla*. Because of very low cone and seed production as well as difficult access for a few populations, only one or two individual trees were sampled from some populations. Also, the sample sizes in a number of populations were relatively small in spite of most cone-bearing trees being sampled. Consequently, cones were collected individually from 98 trees of 15 populations, representing all the four regions and most populations within regions, i.e. all four populations in region BM (28 trees), six populations in region DL (32 trees), four populations in YC (30 trees), and one population in DY (8

trees). The locations and sample sizes of these populations are provided in Fig. 1 and Table 3. The straight-line distances between the sampled populations ranged from less than 1 km to *c.* 700 km between populations from different regions. Seeds from collected cones were extracted by hand, and were stored at –20 °C until needed.

DNA extraction, amplification, and sequencing

One seed was randomly chosen from each tree and the haploid megagametophytes (the maternal tissue surrounding the embryo of seeds) were used for total DNA extraction. Before DNA extraction, seeds were wrapped in paper towels moistened with distilled water for 24 h at room temperature to make fresh tissue preparation easier.

A preliminary screening for mtDNA polymorphism was conducted using 15 primer pairs (Table S1, Supplementary material) over a panel of 8–12 individuals, at least one tree from each population region. Of these primers, eight amplified expected polymerase chain reaction (PCR) products and produced sequences of high quality (Table S1, Supplementary material). However, sequence polymorphism was observed only in two fragments, the mitochondrial NAD dehydrogenase subunit 1 gene (*nad1*) and NAD dehydrogenase subunit 4 gene (*nad4*). One polymorphic site was found in the second intron of *nad1* and two polymorphic sites were found in the first intron and second exon of *nad4*, respectively. Therefore, two mtDNA loci were further used to estimate mtDNA diversity for all *C. argyrophylla* populations. To remove the uninformative region of *nad1* that was about 2900 bp in length, we designed an internal reverse primer (5′-CCAGCGATTC-CTTCATCAAT-3′) and amplified a fragment of 537 bp in length by the new and original forward primer.

After a preliminary survey on 101 nuclear DNA regions (Temesgen *et al*. 2001; Dvornyk *et al*. 2002; Brown *et al*. 2004), we obtained eight loci that generated a single and strong amplification (Table S2, Supplementary material). To evaluate genomic diversity of *C. argyrophylla*, for each locus we sequenced 20 megagametophytes sampled from four regions, five megagametophytes for each region and at least one individual from each population. All the primers for amplifying the eight loci are listed in Table S2 (Supplementary material). Because locus *2009* showed the highest level of intraspecific variation, we further sequenced the remaining trees for this locus. Finally a total of 98 megagametophytes (each from one of the 98 trees) sampled from all 15 populations were sequenced for the *2009* locus.

Amplification reactions were carried out in a volume of 20 μl containing of 1 × PCR buffer, 1.6 mM $MgCl_2$, 0.2 μM each primer, 0.1 mM each dNTP, 0.5 U of ex-*Taq* DNA polymerase, and 10–50 ng of template. Amplification was carried out in Tgradient 96 U thermocycler (Biometre) as follows: 4 min at 94 °C followed by 36 cycles of 30 s at 94 °C, 30 s at 54–57 °C, 90 s at 72 °C and a final extension at 72 °C for 10 min for loci *2009*, *1643*, *pch*, and *comt* and mitochondrial *nad1*. For the remaining fragments (*0624*, *cad*, *glyhmt*, a-*tub*, and *nad4*), a touchdown procedure was employed. After an initial 3-min denaturation at 94 °C, the reaction tubes went through 11 cycles at a 0.5 °C descending series of annealing temperatures each cycle (57 °C–52 °C), and 26 cycles with an annealing temperature of 52 °C, denaturation was at 94 °C (1 min) and extension was at 72 °C (2 min) in all cycles, with a final 10-min extension at 72 °C to end the reaction.

Amplification products were separated by electrophoresis on 1.5% agarose gels stained with ethidium bromide, and gel-purified with a Pharmacia purification kit (Amersham Pharmacia Biotech) or DNA Fragment Quick Purification/Recover Kit (Dingguo). Sequencing reactions were conducted using the DYEnamic ET Terminator Kit (Amersham Pharmacia Biotech), following the manufacturer's protocol. Sequencing was carried out on a MegaBACE 1000 automatic DNA sequencer (Amersham Pharmacia Biotech) after sequencing reaction product was purified through precipitation with 70% ethanol and 0.04 M sodium acetate (pH 5.2). All samples were sequenced in both directions, singleton polymorphism was verified by re-amplifying and sequencing the same megagametophyte. Unique haplotypes of mitochondrial fragments and nuclear alleles are deposited in the GenBank under the accession nos DQ424818–DQ424855.

Data analysis

DNA sequences were aligned with CLUSTAL_X 1.81 (Thompson *et al*. 1997) and refined manually. To measure level of genetic variation, average pairwise differences per base pair between sequences (π) (Nei & Li 1979), Watterson's estimates (θ_w) (Watterson 1975), and haplotype diversity (H_d) were calculated using DNASP version 4.10 (Rozas *et al*. 2003). Tajima's *D* (Tajima 1989) and Fu and Li's *D** (Fu & Li 1993) neutrality tests were used to determine whether a locus is evolving neutrally and is therefore appropriate for a phylogeographical study (Caicedo & Schaal 2004). A phylogenetic network (tree) was constructed by coalescent simulations using the Median-Joining model implemented in NETWORK version 4.0 (Bandelt *et al*. 1999). We used an unrooted network that is a more accurate way to represent haplotype relationships at the intraspecific level than is a bifurcating tree (Posada & Crandall 2001; Caicedo & Schaal 2004). Haplotype network can be used reliably for inference of population history because the most ancient haplotypes should be located at the centre of the gene tree and be geographically widespread whereas the most recent haplotypes should be at the tips of the gene tree and be localized geographically (Schaal *et al*. 1998).

A hierarchical analysis of population subdivision was performed for the four regions and for the populations within regions using an analysis of molecular variance (AMOVA) as implemented in ARLEQUIN 2000 (Schneider *et al*. 2000). AMOVA estimates the relative contribution to molecular variance of differences among groups, differences among populations and differences among populations within groups (Excoffier *et al*. 1992). A comparison was made between G_{ST} and N_{ST} using the *U*-statistic, which is approximated by a Gaussian variable by taking into account the covariance between N_{ST} and G_{ST}, and a one-sided test (Pons & Petit 1996). G_{ST} makes use of haplotype frequencies while N_{ST} takes into account differences between haplotypes. A higher N_{ST} than G_{ST} usually indicates the presence of phylogeographical structure (Pons & Petit 1996) with closely related haplotypes being found more often in the same area than less closely related haplotypes. Test of correlation between geographical and genetic distance between populations were carried out with MANTEL version 2.0 (Liedloff 1999).

Demographical events can leave characteristic signatures in the distribution of pairwise nucleotide differences between individuals, or mismatch distribution (Slatkin & Hudson 1991; Rogers & Harpending 1992). The mismatch distribution is a widely accepted molecular approach to examine genealogical diversity by calculating the number of mutational differences between individual haplotypes within a population (Rogers & Harpending 1992; Johannesen *et al*. 2005). Multimodal mismatch distributions are assumed to characterize old populations of constant size whereas expanding populations are expected to be unimodal (Harpending 1994). We used the mismatch distribution approach to evaluate the demographic history of *C. argyrophylla*. Mismatch distribution for the species as well as for each region were calculated separately with the expected frequency based on a population growth-decline model using DNASP. The ruggedness index (*r*) could be used to quantify the smoothness of mismatch distribution (Harpending 1994). Under the population growth model, the *r* values are expected to be typically low and can be tested for deviation from the constant population size model by simulation implemented in DNASP.

Results

Mitochondrial DNA haplotypes and their distribution

Of eight mtDNA regions with 9860 bp in length surveyed, only two (*nad1* and *nad4*) showed polymorphism in *Cathaya argyrophylla* populations. These two loci contained three substitutions without length variation. Of the three substitutions, one was G/A at site 264 of *nad1* intron 1 and the other two were all A/C at sites 214 and 521 of *nad4*, respectively (Table 1). Two haplotypes were detected for each of *nad1* intron 2 and *nad4* intron 1 and thus a theoretical maximum of haplotypes was four if recombination occurred between the two loci. In total, three distinct mitochondrial haplotypes (mitotypes) were determined by considering simultaneously the information at both mitochondrial loci, i.e. GAA, ACC, and GCC as named after their nucleotide constitutions in polymorphic sites. The designation of the mitotype variants and their geographical distribution are provided in Table 1 and Fig. 2. No single population and region showed polymorphism and the subdivision was therefore maximal ($G_{ST} = N_{ST} = 1$). It is obvious that the distribution of these mitotypes clearly divides the range of *C. argyrophylla* into three: the first mitotype (MT1) was found in all populations of the BM region; the second (MT2) was present in all populations of the DL region, whereas the last mitotype (MT3) was detected in all populations of the YC and DY regions (Fig. 2).

Table 1 Designation of the mitochondrial variants detected at two polymorphic mtDNA loci examined in 15 *Cathaya argyrophylla* populations

	nad1 intron 2	*nad4* intron 1	
Mitotype	264	214	521
MT1	G	A	A
MT2	A	C	C
MT3	G	C	C

Sequence variation at nuclear loci

Twenty megagametophytes (five for each of the four distinct regions and at least one from each population) were sequenced for eight loci, with the amplified fragments ranging from 547 to 808 bp in length. Nucleotide diversity at each locus and in each geographical region is presented in Table 2. The nucleotide diversity varied across loci from 0.0 to 0.0083 (θ_{ws}) and 0.0–0.0092 (π_s) (Table 2). Loci *glyhmt* and *pch* had the lowest diversity ($\theta_{ws} = 0.0$; $\pi_s = 0.0$) while locus *2009* the highest ($\theta_{ws} = 0.0083$; $\pi_s = 0.0092$). The average estimates of nucleotide diversity over the eight loci varied among regions. The DL region had the highest diversity ($\theta_{ws} = 0.0026$; $\pi_s = 0.0028$), followed by DY ($\theta_{ws} = 0.0023$; $\pi_s = 0.0023$) and YC ($\theta_{ws} = 0.0018$; $\pi_s = 0.0017$), and the BM region maintained the lowest ($\theta_{ws} = 0.0012$; $\pi_s = 0.0015$). The species-wide estimate of overall nucleotide diversity was $\theta_{ws} = 0.0024$ and $\pi_s = 0.0029$ (Table 2).

Neutrality tests and phylogeographical analysis of the 2009 *locus*

Of the eight nuclear loci, *2009* showed the highest diversity and thus were informative for further application in

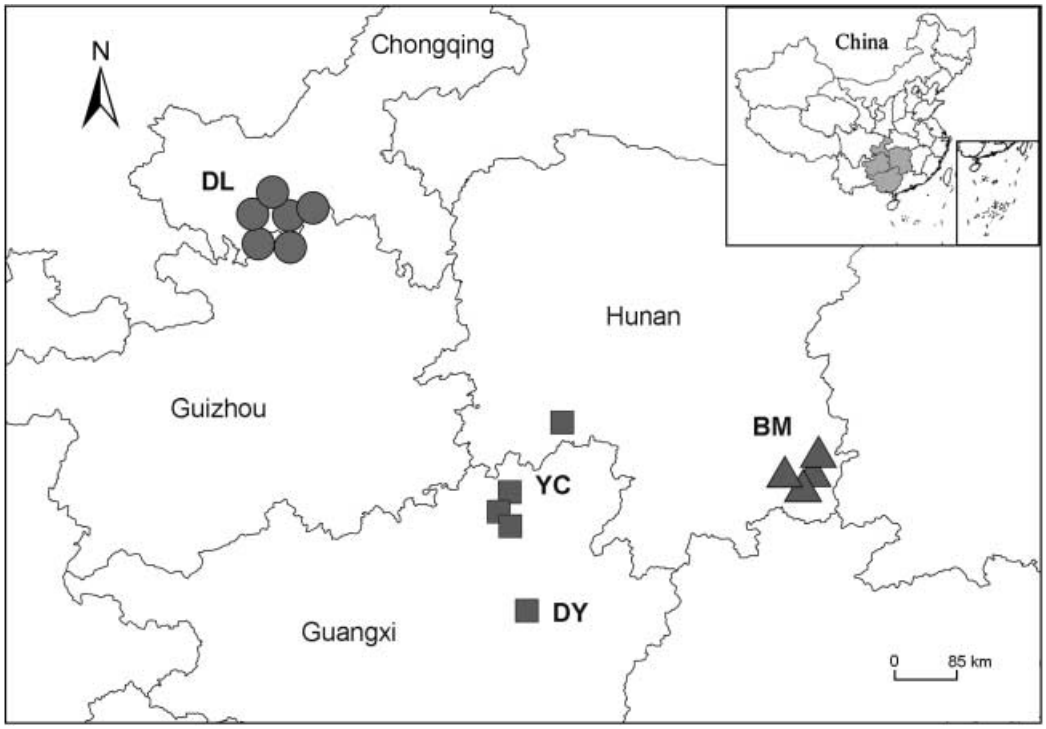

Fig. 2 Distribution of *Cathaya argyrophylla* mitotypes. Definition of the mitotypes was presented in Table 1. ▲, MT1; ●, MT2; ■, MT3.

Table 2 Summary statistics of sequence variation for each population region and the species at the eight nuclear loci

		Locus								
	Region	*2009*	*1643*	*cad*	*0624*	*pch*	*comt*	*glyhmt*	*a-tub*	Average
N		20	20	20	20	20	20	20	20	
L		568	612	808	797	547	715	608	705	5360
S		10	4	10	3	2	3	1	1	34
H_d		0.837	0.563	0.837	0.721	0.416	0.100	0.189	0.337	0.500
θ_w	BM	0.0042	0.000	0.0025	0.0006	0.0000	0.0000	0.0000	0.0000	0.0009
	DL	0.0059	0.0032	0.0037	0.0012	0.0018	0.0000	0.0008	0.0000	0.0021
	DY	0.0059	0.000	0.0025	0.0012	0.0009	0.0020	0.0008	0.0007	0.0018
	YC	0.0025	0.0032	0.0006	0.0018	0.0000	0.0000	0.0000	0.0007	0.0011
	Species	**0.0050**	**0.0019**	**0.0036**	**0.0011**	**0.0010**	**0.0012**	**0.0005**	**0.0004**	**0.0018**
θ_{ws}	BM	0.0063	0.0000	0.0025	0.0008	0.0000	0.0000	0.0000	0.0000	0.0012
	DL	0.0095	0.0060	0.0037	0.0015	0.0000	0.0000	0.0000	0.0000	0.0026
	DY	0.0094	0.0000	0.0025	0.0015	0.0000	0.0038	0.0000	0.0009	0.0023
	YC	0.0047	0.0060	0.0006	0.0023	0.0000	0.0000	0.0000	0.0009	0.0018
	Species	**0.0083**	**0.0035**	**0.0036**	**0.0013**	**0.0000**	**0.0022**	**0.0000**	**0.0005**	**0.0024**
π	BM	0.0053	0.0000	0.0031	0.0005	0.0000	0.0000	0.0000	0.0000	0.0011
	DL	0.0053	0.0040	0.0047	0.0013	0.0018	0.0000	0.0007	0.0000	0.0022
	DY	0.0067	0.0000	0.0021	0.0013	0.0007	0.0017	0.0007	0.0009	0.0018
	YC	0.0028	0.0026	0.0005	0.0017	0.0000	0.0000	0.0000	0.0006	0.0010
	Species	**0.0062**	**0.0035**	**0.0037**	**0.0013**	**0.0008**	**0.0004**	**0.0003**	**0.0005**	**0.0021**
π_s	BM	0.0079	0.0000	0.0031	0.0006	0.0000	0.0000	0.0000	0.0000	0.0015
	DL	0.0085	0.0075	0.0047	0.0016	0.0000	0.0000	0.0000	0.0000	0.0028
	DY	0.0105	0.0000	0.0021	0.0016	0.0000	0.0031	0.0000	0.0010	0.0023
	YC	0.0053	0.0050	0.0005	0.0022	0.0000	0.0000	0.0000	0.0007	0.0017
	Species	**0.0092**	**0.0065**	**0.0037**	**0.0016**	**0.0000**	**0.0008**	**0.0000**	**0.0006**	**0.0029**

N, trees sampled; L, length of the sequenced fragments; S, number of segregating sites; H_d, haplotype diversity; π and π_s, nucleotide diversity (Nei & Li 1979) for the total and silent sites, respectively; θ_w and θ_{ws}, Watterson's parameter (Watterson 1975) for the total and silent sites, respectively.

Table 3 Haplotype distribution and measures of nucleotide diversity at locus *2009*

		Haplotype														
Population	*N*	H1	H2	H3	H4	H5	H6	H7	H8	S	*H*	H_d	θ_w	θ_{ws}	π	π_s
BM1	16	11					5			4	2	0.458	0.0021	0.0030	0.0032	0.0045
BM2	5			1			3		1	2	3	0.700	0.0017	0.0016	0.0018	0.0020
BM3	5	3		2						5	2	0.600	0.0042	0.0063	0.0053	0.0079
BM4	2						2			—	—	—	—	—	—	—
Subtotal	**28**	**14**		**3**			**10**		**1**	**6**	**4**	**0.494**	**0.0023**	**0.0031**	**0.0031**	**0.0043**
DL1	3	1	1	1						7	3	1.000	0.0082	0.0131	0.0082	0.0131
DL2	14	1		4	1	6	2			8	5	0.758	0.0044	0.0072	0.0039	0.0069
DL3	3		1	2						3	2	0.667	0.0035	0.0066	0.0035	0.0066
DL4	2			1			1			—	—	—	—	—	—	—
DL5	1			1						—	—	—	—	—	—	—
DL6	9	1		1		3	4			7	4	0.750	0.0045	0.0073	0.0037	0.0062
Subtotal	**32**	**3**	**2**	**10**	**1**	**9**	**7**			**10**	**6**	**0.771**	**0.0047**	**0.0078**	**0.0042**	**0.0073**
DY	**8**		**1**	**1**			**1**	**5**		**8**	**4**	**0.643**	**0.0054**	**0.0089**	**0.0060**	**0.0095**
YC1	6	3				3				6	2	0.600	0.0046	0.0072	0.0063	0.0099
YC2	6	2		1		2	1			7	4	0.867	0.0054	0.0086	0.0062	0.0099
YC3	12	2		1		6	3			7	4	0.710	0.0041	0.0065	0.0044	0.0071
YC4	6	2		2			2			5	3	0.800	0.0039	0.0058	0.0047	0.0070
Subtotal	**30**	**9**		**4**		**11**	**6**			**7**	**4**	**0.737**	**0.0044**	**0.0069**	**0.0052**	**0.0082**
Total	98	26	3	18	1	20	24	5	1	12	8	0.799	0.0041	0.0064	0.0053	0.0084
(%)		(26.5)	(3.1)	(18.4)	(1.0)	(20.4)	(24.5)	(5.1)	(1.0)							

N, trees sampled; *S*, number of segregating sites; *H*, number of haplotypes; H_d, haplotype diversity; π and π_s, nucleotide diversity (Nei & Li 1979) for the total and silent sites, respectively; θ_w and θ_{ws}, Watterson's parameter (Watterson 1975) for the total and silent sites, respectively.

phylogeographical study. The sequenced *2009* fragment had a length of 568 bp in *C. argyrophylla*, including 333 bp in exon and 235 bp in intron. Based on the sequences of the 98 individuals, 12 polymorphic sites were observed with two singletons and without insertion/deletion (indel). Eight haplotypes were found in the species (Table 3; Fig. S1, Supplementary material). All the four regions had at least four haplotypes, with DL having six and the remaining regions having four (Table 3). Two haplotypes (H3 and H6) were found in 10–12 populations from all four regions (observed over 18 times and more than 18%). H1 was the most frequent haplotype (26.5%) and present in nine populations in three regions except for DY. Haplotype H5 was a frequent haplotype (20.4%) and detected in five populations from two distant regions, DL and YC. All the other haplotypes were observed less than five times (< 5.1%) in no more than two regions. H2 was rare haplotype and found in three populations from the DL and DY regions. Haplotype H7, which is one mutational step away from H1, occurred only in the DY population but was the most frequent haplotype (63%) in this region. Haplotypes H4 and H8 (observed one time) were found only in the DL and BM regions, respectively (Table 3, Fig. 3, Fig. S1, Supplementary material). Haplotype diversity (H_d) was highest for the DL region (0.771), followed by YC (0.737) and DY (0.643) and was the lowest for the BM region (0.494) (Table 3).

To determine if the *2009* locus evolves neutrally, neutrality test was performed for the entire and population datasets separately. Neither Tajima's *D* with the entire dataset ($D = 0.773$, $P > 0.1$) nor Tajima's *D* with the regional datasets ($D = -0.187$–1.958, $P > 0.1$) rejected the null hypothesis of neutral evolution. Fu and Li's D^* test did not rejected the null hypothesis either ($D^* = 0.219$–1.273, $P > 0.1$) (Table 4). Because the *2009* fragment has high level of variation and low recombination and evolve neutrally, variation patterns of this locus might reliably reflect the population history of *C. argyrophylla*. Therefore, haplotype network was used to infer the ancestral (internal) vs. derived (tip) relationships among haplotypes. In the *2009* haplotype network (Fig. 3), four widespread and high frequency haplotypes (H1, H3, H5, and H6) located at interior node and could be considered as ancestral. The rare haplotypes (H2, H4, H7, and H8) were derived from the ancestral haplotypes and were at the tips of the networks (Fig. 3).

Population genetic structure

An AMOVA revealed high amounts of variation both between and within populations at the *2009* locus. Although the

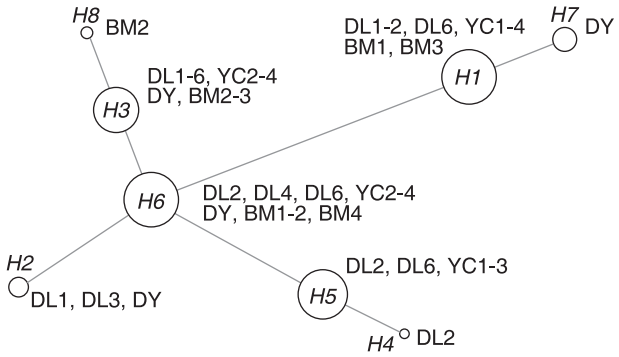

Fig. 3 Haplotype network of the nuclear *2009* fragment. H1-H8 are the haplotypes and correspond to those in Table 3. Population numbers are the same as in Fig. 1 and indicated around the haplotypes. The size of the circle is proportional to the relative frequency of the haplotypes in the study (exact frequencies are shown in Table 3).

majority of nucleotide diversity (77.73%, permutation test $P < 0.001$) was attributable to variation within populations, significant proportions of diversity were attributable to differences among regions (13.40%; permutation test $P < 0.05$) and among populations within regions (8.88%; permutation test $P < 0.05$) (Table 5). F_{ST} analysis also indicated significantly high differentiation among the sampled populations (0.223, $P < 0.001$) and significant pairwise differentiation between all regions (F_{ST} = 0.119–0.377; $P < 0.001$) except for that between DL and YC (F_{ST} = 0.046; $P > 0.05$) (Table 4). To further investigate the population structure within regions, we conducted the F_{ST} analyses for three distinct regions (DL, YC, and BM) where multiple populations were sampled. It is interesting that high level of genetic differentiation among populations were observed for the BM region (F_{ST} = 0.339, $P < 0.001$) (Table 4).

No significant association was detected between genetic and geographical distances (Mantel's test: $r = 0.266$, $P > 0.05$), which do not support the isolation by distance between *C. argyrophylla* populations. A test for phylogeographical structure of haplotype variation across the distribution of the species also demonstrated that the difference between N_{ST} (0.138) and G_{ST} (0.134) was not significant ($P > 0.05$), indicating that closely related haplotypes were not found more often in the same area than less closely related haplotypes.

Mismatch distribution analysis

To test for the hypothesis of population expansion in *C. argyrophylla*, we computed the distribution of pairwise differences from segregation sites of *2009* haplotypes. The DY region was excluded from analysis because of its low sample size. The mismatch distributions for all three regions and for the species were clearly not unimodal and of different shape as that expected for an expanding population (a bell-shape curve) (Fig. 4). The raggedness index (r) of mismatch distribution ranged from 0.0268 to 0.2943 and did not deviate significantly from the constant population sizes for the regions and the species ($P > 0.05$) (Fig. 4). Therefore, these data suggested that neither the BM, DL, and YC regions nor the species underwent population expansion. The positive values of Tajima's D and Fu & Li's D^* statistics for the regions and for the entire sample also implied no evidence of population expansion in *C. argyrophylla* (Table 4).

Table 4 Tajima's D and Fu and Li's D^* tests as well as the F_{ST} between and within regions

	Tajima's D	Fu & Li's D^*	BM	DL	DY	YC
BM	1.615^N	0.457^N	0.339***			
DL	–0.187^N	0.857^N	0.268***	0.009^N		
DY	0.538^N	0.219^N	0.120***	0.377***	—	
YC	1.958^N	1.273^N	0.119***	0.046^N	0.214***	–0.001^N
Total	0.773^N	0.219^N	0.173***			

***$P < 0.001$; ^{N}, not significant.

Table 5 Results from the analyses of molecular variance (AMOVA) for 15 populations of *Cathaya argyrophylla*

Source of variation	d.f.	Sum of squares	Variance components	Variation (%)
Among regions	3	21.976	0.20980	13.40*
Among populations within regions	10	20.493	0.13898	8.88*
Within populations	83	101.016	1.21706	77.73***
Total	96	143.485	1.56584	

d.f., degree of freedom; *$P < 0.05$; ***$P < 0.001$.

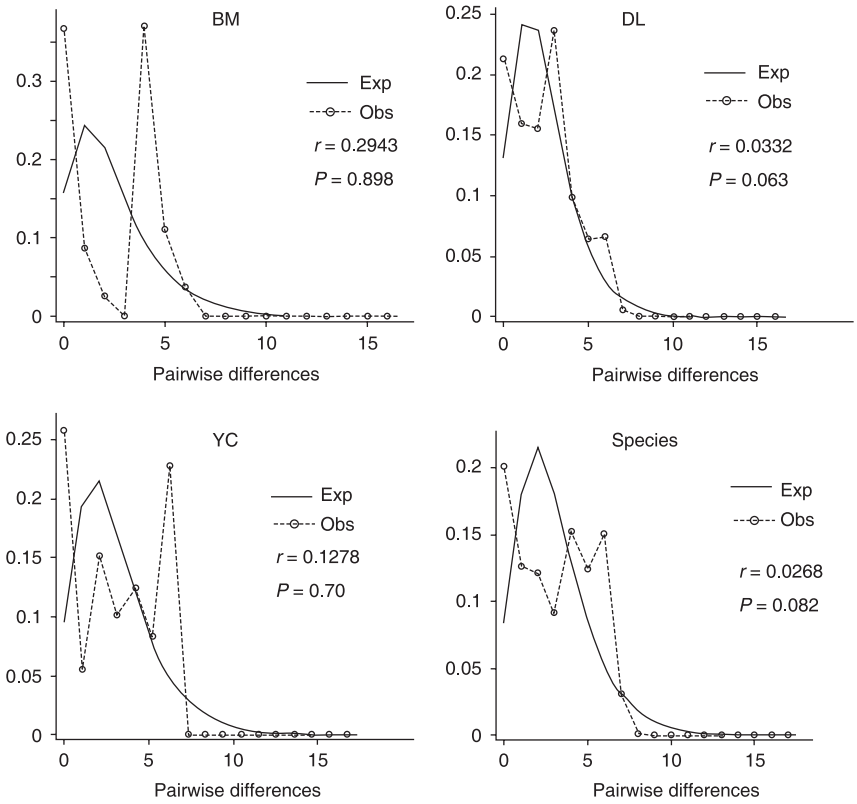

Fig. 4 Mismatch distribution for the BM, DL, YC regions and the species, respectively. The line represents the distributions expected for an expanding population, and the dotted line represents the observed mismatch distribution from segregating sites of the aligned 2009 sequences.

Discussion

Genetic diversity and population structure

In the present study, we found only three substitutions at mtDNA sequences without variation at population and regional levels ($H_E = 0.0$) although eight fragments (a total of 9860 bp) were sequenced. This result is not unexpected because little or no intraspecific variation for mtDNA sequences was previously observed in many conifers including species in the genera *Picea* (Gamache *et al*. 2003; Jaramillo-Correa *et al*. 2004), *Pinus* (Sinclair *et al*. 1999; Soranzo *et al*. 2000; Burban & Petit 2003), and *Larix* (Semerikov & Lascoux 2003; Gros-Louis *et al*. 2005). For the nuclear loci, the mean of silent nucleotide diversity across all eight loci ($\theta_{ws} = 0.0024$; $\pi_s = 0.0029$) are comparable to those found in *Pinus sylvestris* and *Pinus pinaster* but much lower than those found in many other conifers ($\theta_{ws} = 0.0030$–0.0153, $\pi_s = 0.0038$–0.0122) based on estimates of multiple loci (Table 6). However, it should be noted that the nucleotide diversity estimated for *P. sylvestris* was based only on two loci (Garcia-Gil *et al*. 2003) and those for *P. pinaster* and *Pinus radiata* were based on genes involved in wood formation and were possibly under strong selection (Pot *et al*. 2005). Therefore, the species-wide nucleotide diversity in *Cathaya argyrophylla* is the lowest among the conifers for which multiple genes have been sequenced (Table 6). At regional level, the lowest (about one half of the total) diversity was found in the BM region ($\theta_{ws} = 0.0012$; $\pi_s = 0.0015$), followed by YC ($\theta_{ws} = 0.0018$; $\pi_s = 0.0017$) and relatively higher diversity occurred in the other two regions ($\theta_{ws} = 0.0023$–0.0026; $\pi_s = 0.0023$–0.0028) (Table 2). In the *2009* data set with a larger sample, some populations (i.e. BM1 and BM2) harboured much less variation than the others. Also, the lowest diversity was observed in the BM region, although as many as 28 trees from four populations were sampled from this region (Table 3).

In addition to low nucleotide diversity, another important feature for *C. argyrophylla* is that genetic differentiation among populations and regions is significantly higher than those found in other conifers. By comparison with other conifers for the maternally inherited markers (mean $G_{ST} = 0.76$, averaged over 33 conifers; Petit *et al*. 2005), genetic differentiation between regions at mtDNA markers is extremely large ($G_{ST} = 1.0$), indicating particularly low seed flow in the species. Similar patterns of genetic structure have been documented in many conifers in Europe and North America (e.g. Sinclair *et al*. 1999; Mitton *et al*. 2000; Richardson *et al*. 2002; Burban & Petit 2003; Jaramillo-Correa *et al*. 2004). In addition to the complete fixation of three mitochondrial haplotypes in different regions, about 22% of the total diversity at the nuclear *2009* locus is attributed to variation among populations in *C. argyrophylla*. This value is significantly higher than the

Table 6 Estimates of nucleotide diversity in different coniferous species

Species	No. of loci	Length (bp)	θ_{sil}	π_{sil}	Species-wide	Reference
Cathaya argyrophylla	8	5360	0.0024	0.0029	Yes	This study
Cryptomeria japonica	7	10158	0.0030	0.0038	Yes	Kado *et al*. (2003)
Pinus densata	7	3040	0.0143	0.0122	Yes	Ma et al. (2006)
Pinus pinaster	10	4746	0.0017*	0.0024*	Yes	Pot *et al*. (2005)
Pinus radiata	10	4746	0.0008*	0.0007*	No	Pot *et al*. (2005)
Pinus tabuliformis	7	3040	0.0153	0.0119	Yes	Ma *et al*. (2006)
Pinus sylvestris	2	4136	0.0030	0.0019	Yes	Garcia-Gil *et al*. (2003)
Pinus taeda	19	17580	0.0066	0.0064	Yes	Brown *et al*. (2004)
Pinus taeda	18	10116	0.0079	0.0085	No	Gonzalez-Martinez *et al*. (2006)
Pinus yunnanensis	7	3040	0.0077	0.0095	Yes	Ma *et al*. (2006)
Pseudotsuga menziesii	18	15183	0.0113	0.0106	No	Krutovsky & Neale (2005)

θ_{sil}, average, θ_W at silent sites; π_{sil}, average π at silent sites; *noncoding region.

average (G_{ST} = 0.073) of allozyme data from more than 100 gymnosperms (Hamrick *et al*. 1992) and the average (G_{ST} = 0.116) summarized recently by Petit *et al*. (2005) based on biparentally inherited markers of 33 conifers. It is worthwhile mentioning that the genetic differentiation among populations within the BM region is as high as 34% (Table 4) even though these four populations were sampled in an area of less than approximately 2×3 km^2.

Based on allozyme study on eight *C. argyrophylla* populations sampled from the same four regions, Ge *et al*. (1998) detected a low amount of variation (P = 30.4%; H_E = 0.102) and five times higher level of interpopulation differentiation (G_{ST} = 0.441) compared to the average (P = 53.4%; H_E = 0.151; G_{ST} = 0.073) of more than 100 gymnosperms (Hamrick *et al*. 1992). Our present results based on sequence data corroborates our previous findings. First, habitat fragmentation and deterioration resulting from Quaternary glaciation are obviously the most likely explanations for the disjunct distribution and small population sizes of *C. argyrophylla* (Liu & Basinger 2000), which in turn lead to low genetic diversity and high population differentiation in the species. Habitat fragmentation is supported by the finding that four alleles (haplotypes) are present in the widely separated regions (Table 3, Fig. 3, and Fig. S1, Supplementary material), in agreement with the fossil evidence that indicated fragmentation of *C. argyrophylla* populations during the Quaternary (Liu & Basinger 2000). Theoretically, reductions in population size create genetic bottlenecks because remaining individuals contain only a small sample of the alleles present in the parental generation (Ellstrand & Elam 1993). This could account for the particularly low levels of diversity observed in some populations and regions, especially in the BM region that is located in the most eastern part of current distribution of the species. Much higher value of F_{ST} compared to the other conifers and no correlation between genetic distance and geographical distance, even at regional level (Table 4) are indicative of genetic drift in this species. Genetic drift has also been documented in our previous allozyme data where the same alleles at many loci were randomly fixed in geographically distant populations (Ge *et al*. 1998). Our phylogeographical study further suggests that the population contractions and bottlenecks have happened relatively recently. As evidenced in a study of *Pinus resinosa* in North America (Echt *et al*. 1998), a single, common, and often frequent, haplotype would be detected if a major ancient population bottleneck event had happened for the species.

Second, limited gene flow would be another factor for the marked genetic differentiation among populations in *C. argyrophylla*. It is reported that the pollination and mating system of *C. argyrophylla* is little different from other conifers, and the primary seed dispersal is through gravity and wind, with secondary dispersal facilitated by squirrels (Wang 1990; Xie & Chen 1999). However, the squirrel dispersal did not work effectively enough, partly because of very low rates of seed germination and seedling survival (Xie & Chen 1999). The estimated value of *Nm* (1.85 immigrants per generation) in *C. argyrophylla* is lower relative to those (*Nm* = 3.4–17.2) in pines in range-wide studies (reviewed in Ledig 1998). Such small amounts of gene flow in *C. argyrophylla* are not sufficient to counteract genetic drift in the absence of selection (Ellstrand & Elam 1993). In particular, much of the gene flow measured by *Nm* is probably the reflection of ancient gene flow, as evidenced by wide distribution of the ancestral haplotypes at the *2009* loci (H1, H3, H5, and H6). Therefore, recent gene flow by pollen would be greatly reduced, as evidenced by significantly high genetic differentiation of nuclear markers among populations (F_{ST} = 0.223) and among regions (F_{ST} = 0.173) (Table 4). It should be mentioned that although great efforts have been made, the sample sizes per population in the present study were relatively small, which may lead to approximate estimates of some statistics.

Demographic history and potential refugia

The late Quaternary palaeoecology of North America and Europe has revealed a series of southward range contractions during the latest glacial period followed by rapid northward range expansions following deglaciation (Taberlet *et al*. 1998; Mitton *et al*. 2000; Petit *et al*. 2003; Hewitt 2004). Although relatively few studies have been conducted on the postglacial history of Chinese plants, phylogeographical studies on a number of endemic species in southern China have revealed postglacial population growth. Based on a phylogeographical analysis of a conifer (*Cunninghamia konishii*), Lu *et al*. (2001) and Hwang *et al*. (2003) inferred several refugia in southern China and detected postglacial population growth after a glacial bottleneck event. Shen *et al*. (2005) also found possible refugia of *Ginkgo biloba* in southwestern China and did not support previous consideration that the Tianmu Mountain in eastern China was a potential refugium for the species.

Unlike most conifers mentioned above, we do not find evidence of long distance dispersal and population expansion in *C. argyrophylla*. As indicated by Printzen *et al*. (2003), large-scale infraspecific disjunctions in many species could be explained alternatively by range fragmentation and widespread long distance dispersal. In our case, the presence of nonoverlapping mitotypes and comparable levels of nuclear nucleotide diversity in different regions make the dispersal scenario unlikely, because the descendant populations would harbour much reduced genetic diversity relative to those in the refugia and such a reduction in diversity would increase with distance from a refugium (Comes & Kadereit 1998; Schaal *et al*. 1998; Petit *et al*. 2003). As is evident from Fig. 3, all the widespread haplotypes (H1, H3, H5, and H6) represent internal nodes of the network and would predate the divergence of the populations (ancient). Their distribution across all the regions probably owes more to the persistence of ancestral polymorphisms than to recent gene flow and dispersal (Schaal *et al*. 1998). In addition, the pattern of haplotype distribution is not consistent with a rapid range expansion where a 'star-like' phylogeny of haplotypes would be expected, as evidenced in many other conifers (Echt *et al*. 1998; Hwang *et al*. 2003). The lack of population expansion was further supported by the multimodal mismatch distribution (Fig. 4), and the positive Tajima's *D* and Fu & Li's *D** values (Table 4).

Hewitt (2000) indicated that in southern temperate regions and the tropics, the varied topography tends to subdivide the species into populations that may evolve independently with only occasional gene flow, perhaps only every glaciation or interstadial. This is probably the case in *C. argyrophylla*. Based on extensive investigations on *Cathaya* pollen and tabulate data from other extant coniferous genera with bisaccate pollen, Liu & Basinger (2000) outlined the palaeogeographical history of the genus and indicated that *Cathaya* was apparently restricted to North America and East Asia during the Cretaceous and became widespread in North America, East Asia, and Europe in the Neogene. They suggested that Late Tertiary climatic deterioration and Quaternary glaciation would have been responsible for expiration of *Cathaya* from North America first and then from Europe and thus the endemic distribution of extant *Cathaya* in China represents a remnant of a formerly widespread Asiatic population. The present distribution of three mitotypes and the four geographical groups suggest the existence of at least four separate refugia during the last ice age when once widespread species became fragmented during the Quaternary. The refugia include the Dalou Moutains (DL) in the western, the Bamian Moutains (BM) in the Eastern, the Yuechengling Moutains (YC) in the Southeastern, and the Dayao Mountains (DY) in the South (Figs 1 and 2). Although YC and DY share the same mitotype (MT3), they are most likely to represent different refugia because DY, separated from YC by 180 km, has comparable diversity to YC (Table 2) and is genetically isolated from the YC populations ($F_{ST} = 0.214$, $P < 0.001$, Table 4). The presence of a private haplotype (H7) at high frequency in DY (Table 3) also corroborate this suggestion. Multiple refugia have been documented in many other conifers (Konnert & Bergmann 1995; Sinclair *et al*. 1999; Mitton *et al*. 2000; Richardson *et al*. 2002; Burban & Petit 2003; Jaramillo-Correa *et al*. 2004; Godbaut *et al*. 2005) as well as plant species in China (Lu *et al*. 2001; Ge *et al*. 2002; Cheng *et al*. 2005; Shen *et al*. 2005).

Ying *et al*. (1993) and Ying (2001) localized three regions with high levels of plant diversity and endemism in China, i.e. the Hengduan range, the Central China and the Lingnan region. The current *Cathaya* populations are located in the Central China and the Lingnan region. Both regions are characterized by the great diversity in topography, climate, and ecological conditions and spared from the direct effect of the repeated Pleistocene continental glaciation (Hu 1980; Liu & Basinger 2000). Many ancient plant species including the gymnosperm endemics, *Glyptostrubus pensilis*, *Metasequoia glyptostroboides*, *Pseudotaxus chienii*, *Taiwania flousiana*, and *Ginkgo biloba* were also found in these regions (Hu 1980; Ying *et al*. 1993; Shen *et al*. 2005). Therefore, these three regions with high diversity and endemism are most likely to be refugia for plant species in general.

Implications for conservation

The genetic profile uncovered in this study has not only provided important insights into the evolutionary history of *C. argyrophylla*, but is also critical for its conservation management. Although all the *C. argyrophylla* populations in the four distinct regions are currently under legal protection as one of the most endangered plant species in China (Fu 1992; Wang & Xie 2004), further measures

await serious consideration in addition to the routine practices (Ge *et al*. 1998; Xie & Chen 1999). Low level of genetic diversity and high genetic differentiation resulting from population bottleneck and limited gene flow should be taken into consideration in the conservation decision and management of this endangered species, as proposed in our previous study (Ge *et al*. 1998). The significance of the unique population genetic profile and demographic history found in this study is twofold.

First, almost all extant populations in *C. argyrophylla* face serious threat and risk of extinction by stochastic processes because of low genetic diversity and small population sizes. Our field surveys indicated that most populations of this species were small in sizes with a few young trees and only one or two mature individuals in some populations in the BM and DL regions. Population size is the most influential of the five criteria for listing species as endangered under the International Union for the Conservation of Nature and Natural Resources (IUCN) system (Frankham *et al*. 2002). More importantly, the ability of this species to compete with other species and recolonize new habitats is very low as the habitats suitable for this species were reported to be in the process of deterioration and fragmentation (Fu 1992; Xie *et al*. 1999). Unusually low fertility and survival rates were additional factors that limit the population expansion of this species (Sun *et al*. 1994; Xie & Chen 1999). Our field survey and ecological studies have indicated that the majority of the *C. argyrophylla* populations have been already restricted to the cool, moist forests, predominantly on the open slopes and the highest peaks of the mountains in the distinct regions (Xie & Chen 1999). It seems unlikely for the *C. argyrophylla* populations to expand their range without human intervention. The routine stand management can do nothing to conserve and recover the species because the main threats to *C. argyrophylla* are habitat deterioration and loss that might result mainly from global warming (Ledig *et al*. 2002). Therefore, in addition to the *in situ* strategy that is taken currently, *ex situ* conservation should be given high priority to offset the habitat deterioration and fragmentation. In this regard, reintroductions and introductions can be designed to re-establish self-sustaining wild populations and this practice should be carried out into suitable habitats within the previous range of the species (Frankham *et al*. 2002).

Second, a promising measure that is likely to restore and enrich genetically threatened or declining populations is to employ traditional breeding programs such as controlled crossing between genetically distinct populations, i.e. between the four distinct regions. Introgression and hybridization are important genetic management actions for wild populations and have been successfully applied in many species (Frankham *et al*. 2002). In *Scabiosa columbaria*, for example, crosses between populations had fitness 2.5 times that of the crosses within populations (van Treurea *et al*. 1993). Our preliminary investigation on pollen viability of the *C. argyrophylla* trees in the DY and YC populations showed that most mature trees generated normal pollens without significant pollen mortality (unpublished data). Given the high genetic differentiation and extremely low gene flow among regions and populations in *C. argyrophylla*, artificial crossing of individuals between populations or regions and subsequent introduction of the offspring into suitable habitats would be potentially effective measures to alleviate the effects of population fragmentation. However, it should be noted that the potential benefits of crossing genetically divergent populations may be counteracted by outbreeding depression and there are very few cases where it is being used successfully in practice (Frankham *et al*. 2002). Previous ecological studies have shown that *C. argyrophylla* populations in different localities shared similar ecological requirements without significant local adaptation (Wang 1990; Xie & Chen 1999) despite high genetic divergence among them. Therefore, controlled crossing between populations seems reasonable but the practice of crossing populations should be evaluated prior to full implementation.

Acknowledgements

We thank Xiao-Ru Wang and Xiao-Fei Ma for providing primers; Zheng-Yu Liu, Can-Ming Zhang, De-Yuan Hong, Yi-Bo Luo, Hai-Ming Tan, Guo-Fang Zhang, Kun Qin, Ding-Heng Zhang, and Hong Liu for help in field survey and seed collection. We are also grateful to Zhi-Yong Zhang, Fu-Min Zhang and Lin-Bin Zhang, Zhong-Qiang Xie for helpful suggestions and discussion, and to Rui Wang for help in drawing. This study was supported by the National Natural Science Foundation of China (30121003) and the State Key Basic Research and Development Plan of China (G2000046805).

Supplementary material

The supplementary material is available from
http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC3086/MEC3086sm.htm

Table S1 Target region, annealing temperatures and expected sizes of PCR products for primer pairs used to amplify mtDNA regions

Table S2 Summary of the genes, primers and amplification conditions

References

Abbott RJ, Smith LC, Milne RI *et al*. (2000) Molecular analysis of plant migration and refugia in the Arctic. *Science*, **289**, 1343–1346.

Avise JC (2000) *Phylogeography, The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Axelrod DI, Al-Shehbaz I, Raven PH (1996) History of the modern flora of China. In: *Floristic Characteristics and Diversity of East Asian Plants* (eds Zhang A, Wu S), pp. 43–55. Springer, New York.

Bandelt H, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004) Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proceedings of the National Academy of Sciences, USA*, **101**, 15255–15260.

Burban C, Petit RJ (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.

Caicedo AL, Schaal BA (2004) Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology*, **13**, 1871–1882.

Cheng YP, Hwang SY, Lin TP (2005) Potential refugia in Taiwan revealed by the phylogeographical study of *Castanopsis carlesii* Hayata (Fagaceae). *Molecular Ecology*, **14**, 2075–2085.

Chun WY, Kuang KZ (1958) A new genus of Pinaceae, *Cathaya* Chun et Kuang, General Nov, from southern and western China. *Bot Zhurn.*, **43**, 461–470 (in Russian).

Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, **3**, 432–438.

Dvornyk V, Sirvio A, Mikkonen M, Savoleinen O (2002) Low nucleotide diversity at the *pal1* locus in the widely distributed *Pinus sylvestris*. *Molecular Biology and Evolution*, **19**, 179–188.

Echt CS, Deverno LL, Anzidei M, Vendramin GG (1998) Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. *Molecular Ecology*, **7**, 307–316.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.

Fu LK (1992) *China Plant Red Data Book: Rare and Endangered Plants*, Vol. 1. Science Press, Beijing and New York.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Gamache I, Jaramillo-Correa JP, Payette S, Bousquet J (2003) Diverging patterns of mitochondrial and nuclear DNA diversity in sub-Arctic black spruce: imprint of a founder effect associated with postglacial colonization. *Molecular Ecology*, **12**, 891–901.

Garcia-Gil MR, Mikkonen M, Savolainen O (2003) Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Molecular Ecology*, **12**, 1195–1206.

Ge S, Hong DY, Wang HQ, Liu ZY, Zhang CM (1998) Population genetic structure and conservation of an endangered conifer, *Cathaya argyrophylla* (Pinaceae). *International Journal of Plant Sciences*, **159**, 351–357.

Ge XJ, Chiang YC, Chou CH, Chiang TY (2002) Nested clade analysis of *Dunnia sinensis* (Rubiaceae), a monotypic genus from China based on organelle DNA sequences. *Conservation Genetics*, **3**, 351–362.

Godbaut J, Jaramillo-Corea JP, Beaulieu J, Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, **14**, 3497–3512.

Gonzalez-Martinez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB (2006) DNA sequence variation and selection of tag SNPs at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics*, **172**, 1915–1926.

Gros-Louis MC, Bousquet JP, Isabel N (2005) Species-diagnostic markers in *Larix* spp. based on RAPDs and nuclear, cpDNA, and mtDNA gene sequences, and their phylogenetic implications. *Tree Genetics & Genomes*, **1**, 50–63.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests*, **6**, 95–124.

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial-DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Hewitt GM (2000) The genetic legacy of the quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosphical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 183–195.

Hu SY (1980) The Metasequoia flora and its phytogeographic significance. *Journal of the Arnold Arboretum*, **61**, 41–94.

Hu YS, Wang FH (1984) Anatomical studies of *Cathaya* (Pinaceae). *American Journal of Botany*, **71**, 727–735.

Hwang SY, Lin TP, Ma CS *et al.* (2003) Postglacial population growth of *Cunninghamia konishii* (Cupressaceae) inferred from phylogeographical and mismatch analysis of chloroplast DNA variation. *Molecular Ecology*, **12**, 2689–2695.

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Molecular Ecology*, **13**, 2735–2747.

Johannesen J, Lubin Y, Laufs T, Seitz A (2005) Dispersal history of a spider (*Stegodyphus lineatus*) across contiguous deserts: vicariance and range expansion. *Biological Journal of the Linnean Society*, **84**, 739–754.

Kado T, Yoshimaru H, Tsumura Y, Tachida H (2003) DNA variation in a conifer, *Cryptomeria japonica* (Cupressaceae *sensu lato*). *Genetics*, **164**, 1547–1559.

Konnert M, Bergmann F (1995) The geographical distribution of genetic variation of silver fir (*Abies alba*, Pinaceae) in relation to its migration history. *Plant Systematics and Evolution*, **7**, 19–30.

Krutovsky KV, Neale DB (2005) Nucleotide diversity and linkage disequilibrium in cold-hardiness- and wood quality-related candidate genes in Douglas fir. *Genetics*, **171**, 2029–2041.

Ledig FT (1998) Genetic variation in *Pinus*. In: *Ecology and Biogeography of Pinus* (ed. Richardson DM), pp. 251–280. Cambridge University Press, Cambridge, UK.

Ledig FT, Hodgskiss PD, Jacob-Cervantes V (2002) Genetic diversity, mating system, and conservation of a Mexican subalpine relict, *Picea mexicana* Martinez. *Conservation Genetics*, **3**, 113–122.

Liedloff A (1999) *MANTEL version 2.0, Nonparametric Test Calculator*. Queensland University of Technology, Australia.

Liu YS, Basinger LF (2000) Fossil *Cathaya* (Pinaceae) pollen from the Canadian High Arctic. *International Journal of Plant Sciences*, **161**, 829–847.

Lu SY, Peng CI, Cheng YP, Hong KH, Chinag TY (2001) Chloroplast DNA phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of Taiwan. *Genome*, **44**, 797–807.

Ma XF, Szmidt AE, Wang XR (2006) Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide variation at seven gene loci. *Molecular Biology and Evolution*, **23**, 807–816.

Mitton JB, Kreiser BR, Latta RG (2000) Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Molecular Ecology*, **9**, 91–97.

Nei M, Li WH (1979) Mathematical model for studying genetic

variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.

Petit RJ, Aguinagalde I, de Beaulieu JL *et al*. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Petit RJ, Duminll J, Fineschi S *et al*. (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, **16**, 37–45.

Pot D, McMillan L, Echt C *et al*. (2005) Nucleotide variation in genes involved in wood formation in two pine species. *New Phytologist*, **167**, 101–112.

Printzen C, Ekman S, Tonsberg T (2003) Phylogeography of *Cavernularia hultenii*: evidence of slow genetic drift in a widely disjunct lichen. *Molecular Ecology*, **12**, 1473–1486.

Richardson BA, Brunsfeld SJ, Klopfenstein NB (2002) DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology*, **11**, 215–227.

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rozas J, Sanchez-DeI, Barrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Schaal BA, Hayworth DA, Oseni KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**, 465–474.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, *version 2.0: a Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Schonswetter P, Stehlik I, Holderegger R, Tribsce A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547–3555.

Semerikov VL, Lascoux M (2003) Nuclear and cytoplasmic variation within and between Eurasian *Larix* (Pinaceae) species. *American Journal of Botany*, **90**, 1113–1123.

Shen L, Chen XY, Zhang X, Li YY, Fu CX, Qiu YX (2005) Genetic variation of *Ginkgo biloba* L. (Ginkgoaceae) based on cpDNA PCR-RFLPs: inference of glacial refugia. *Heredity*, **94**, 396–401.

Sinclair WT, Morman JD, Ennos RA (1999) The postglacial history of Scots pine (*Pinus sylvestris*) in western Europe: evidence from mitochondrial DNA variation. *Molecular Ecology*, **8**, 83–88.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Soltis DE, Gitzendanner MA, Strenge DD, Solits PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.

Soranzo N, Alia R, Provan J, Powell W (2000) Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. *Molecular Ecology*, **9**, 1205–1211.

Sun M, Tang SH, Chen AH, Yang LY, Wang BN (1994) Studies on the growth of mature embryo of *Cathaya argyrophylla* in vitro and inducing of callus. *Journal of Southwest Forestry College*, **14**, 155–158.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Temesgen B, Brown GR, Harry DE *et al*. (2001) Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics*, **102**, 664–675.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL WINDOWS interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

van Treurea R, Bijlsma R, Ourborg NJ, van Delden W (1993) The significance of genetic erosion in the process of extinction. IV. Inbreeding depression and heterosis effects caused by selfing and outcrossing in *Scabiosa columbaria*. *Evolution*, **47**, 1669–1680.

Wang FX (1990) *The Biology of Cathaya argyrophylla*. Science Press, Beijing, China.

Wang S, Xie Y (2004) *China Species Red List*, Vol. 1. Higher Education Press, Beijing.

Wang XQ, Zou YP, Zhang DM, Hong DY, Liu ZY (1997) Genetic diversity of *Cathaya argyrophylla* with RAPD analysis. *Science in China (Serial C)*, **26**, 436–441.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7** (256), 276.

Xie ZQ (1995) *Cathaya argyrophylla*, an endemic species in China, and related studies. *Chinese Biodiversity*, **3**, 99–103.

Xie ZQ, Chen WL (1999) The endangering causes and preserving strategies for *Cathaya argyrophylla*, a plant endemic to China. *Acta Phytoecologica Sinica*, **23**, 1–7.

Xie ZQ, Chen WL, Lu P, Hu D (1999) The demography and age structure of the endangered plant population of *Cathaya argyrophylla*. *Acta Ecologica Sinica*, **19**, 523–528.

Ying TS (2001) Species diversity and distribution pattern of seed plants in China. *Biodiversity Science*, **9**, 393–398.

Ying TS, Zhang YL, Boufford DE (1993) *The Endemic Genera of Chinese Seed Plant*. Science Press, Beijing.

Zhang Q, Chiang TY, George M, Liu JQ, Abbott RJ (2005) Phylogeography of the Qinghai-Tibetan Plateau endemic *Juniperus przewalskii* (Cupressaceae) inferred from chloroplast DNA sequence variation. *Molecular Ecology*, **14**, 3513–3524.

This work represents part of Hong-Wei Wang's PhD dissertation that deals with the population genetics, evolutionary history and conservation management of the endangered *Cathaya argyrophylla*. Song Ge is a research professor at the State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences. His research focuses on the conservation genetics of rare and endangered plant species in China. His group is also interested in the study of phylogeny, speciation and molecular evolution of rice genus and its relatives.
