Tree Genetics & Genomes (2017) 13:91
DOI 10.1007/s11295-017-1172-6

ORIGINAL ARTICLE

Genetic diversity and population structure in the narrow endemic Chinese walnut *Juglans hopeiensis* Hu: implications for conservation

Yiheng Hu[1] · **Meng Dang**[1] · **Xiaojia Feng**[1] · **Keith Woeste**[2] · **Peng Zhao**[1]

Received: 22 February 2017 / Revised: 2 June 2017 / Accepted: 4 July 2017

Abstract The conservation of narrow endemic species relies on accurate information regarding their population structure. *Juglans hopeiensis* Hu (Ma walnut), found only in Hebei province, Beijing, and Tianjin, China, is a threatened tree species valued commercially for its nut and wood. Sequences of two maternally inherited mitochondrial markers and two maternally inherited chloroplast intergenic spacers, three nuclear DNA sequences, and allele sizes from 11 microsatellites were obtained from 108 individuals of *J. hopeiensis*, *Juglans regia*, and *Juglans mandshurica*. Haplotype networks were constructed using NETWORK. Genetic diversity, population differentiation, and analysis of molecular variance (AMOVA) were used to determine genetic structure. MEGA was used to construct phylogenetic trees. Genetic diversity of *J. hopeiensis* was moderate based on nuclear DNA, but low based on uniparentally inherited mitochondrial DNA and chloroplast DNA. Haplotype networks showed that *J. hopeiensis* haplotypes were different than haplotypes found in *J. regia* and *J. mandshurica*. Allelic variants in nuclear genes that were shared among *J. hopeiensis* populations were not found in *J. regia* or *J. mandshurica*. Sampled populations of *J. hopeiensis* showed clear genetic structure. The maximum parsimony (MP) tree showed *J. hopeiensis* to be distinct from *J. mandshurica* but threatened by hybridization with *J. regia* and *J. mandshurica*. *J. hopeiensis* populations are strongly differentiated from sympatric *Juglans* species, but they are threatened by small population sizes and hybridization.

Keywords Chinese walnut · Hybridization · Conservation · Genetic differentiation · Microsatellites · *Juglans regia* · *Juglans mandshurica*

Communicated by A. M. Dandekar

Electronic supplementary material The online version of this article (doi:10.1007/s11295-017-1172-6) contains supplementary material, which is available to authorized users.

✉ Keith Woeste
kwoeste@fs.fed.us

✉ Peng Zhao
pengzhao@nwu.edu.cn

Yiheng Hu
yihenghu@yeah.net

Meng Dang
a1193371050@qq.com

Xiaojia Feng
332146704@qq.com

[1] Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Taibaibei Road 229, Xi'an, Shaanxi Province 710069, China

[2] USDA Forest Service Hardwood Tree Improvement and Regeneration Center (HTIRC), Department of Forestry and Natural Resources, Purdue University, 715 West State Street, West Lafayette, IN 47907, USA

Published online: 01 August 2017

Introduction

A fundamental issue in ecology and conservation biology is how evolutionary processes influence genetic variation across the whole geographic range of a species (Pouget et al. 2013; Budd et al. 2015). The phylogeography of species is shaped by factors such as evolutionary integrity (Moritz 2002; Broadhurst et al. 2008), demographic expansion, environmental variation, and natural catastrophes (Ruan et al. 2013; Poudel et al. 2014). In addition, patterns of plant distribution and genetic diversity have been affected by human manipulation of the environment since the beginning of the Holocene (Gunn et al. 2010; Poudel et al. 2014). Other human activities, such as the commercial circulation and/or translocation of seeds or seedlings with little or no regard for their provenance, have serious implications for local forest genetic resources (Pollegioni et al. 2011, 2015).

The Chinese walnut (*Juglans hopeiensis* Hu; in Chinese "Ma walnut") is one of four East Asian species in *Juglans* section *Cardiocaryon*, which also includes *Juglans mandshurica*, *Juglans cathayensis*, and *Juglans ailantifolia.* Like all *Juglans*, *J. hopeiensis* is a monoecious, wind-pollinated, temperate deciduous tree (Manning 1978). It grows as a sporadic, rare, endangered endemic tree, narrowly distributed in northern China in the hilly, mid-elevation area near Beijing and Tianjin, including parts of Hebei province (Lu et al. 1999; Chen and Gilbert 2006; Aradhya et al. 2007; Zhao et al. 2014; Hu et al. 2015). The Manchurian walnut (*J. mandshurica*) grows in northern and northeastern China, Korea, Japan, and the far eastern section of Russia (Lu 1982; Bai et al. 2010). Persian walnut or common walnut (*Juglans regia*) grows in wide geographical range, including Eurasia, from China to Western Europe, and Eastern Asia (Manning 1978; Pollegioni et al. 2015). Ma walnut (*J. hopeiensis*) is sympatric with *J. mandshurica* (Zhang and Shao 2015; Xi 1990), which has high-quality wood that has been used in military products, paneling, and furniture, but now it is most familiar as the source of a hard-shelled, ornate nut sold in China as a curio or talisman (Hao et al. 2013). The current demographic decline of the natural population of Ma walnut is attributed to human activities including destruction of its preferred habitats and global climate change (Pei et al. 2006; Hu et al. 2015). Previous research on this walnut species was focused on molecular phylogeny (Aradhya et al. 2007), traditional tree breeding, or descriptions of germplasm collections (Hao et al. 2013; He et al. 2015). In general, the species has not been given enough attention and research (Pei et al. 2006). The population genetics of Ma walnut (*J. hopeiensis*) are presented here for the first time, although photosynthesis (Wang et al. 2005), microsporogenensis (Mu et al. 1990), and germplasm and cultivar relationships (Hao et al. 2007) were previously reported.

J. hopeiensis (Ma walnut), and *J. mandshurica* (Manchurian walnut) are native to China, along with *J. cathayensis* and *Juglans sigillata* (Manning 1978; Fjellstrom and Parfitt 1995; Aradhya et al. 2007). *J. regia* is planted as a crop in Asia, Australia, southern Europe, North Africa, North America, and South America, but its origins remain obscure, so whether it is native to China or an ancient introduction into China is not known (McGranahan and Leslie, 1991; Pollegioni et al. 2015, 2017). Whatever their histories, *J. regia* and *J. mandshurica* are now sympatric with *J. hopeiensis* (Lu et al. 1999; Hu et al. 2017).

The relationship of Ma walnut to other members of the section *Cardiocaryon*, especially *J. mandshurica*, is disputed (Lu et al. 1999; Aradhya et al. 2007). Evidence from randomly amplified polymorphic DNA (RAPD) markers, isozymes, and karyotype analysis indicated this species might have arisen from the recent (from 3.51 to 7.91 Ma ago based on whole-chloroplast genome data; Hu et al. 2017) hybridization of *J. regia* and *J. mandshurica* (Wu et al. 1999; Mu et al. 1990; Hu et al. 2017). Others have suggested that *J. hopeiensis* is a variant of *J. mandshurica* based on the anther characteristics and morphology (Lu et al. 1999). The most comprehensive phylogenetic study concluded that Ma walnut is a well-defined lineage and a sister clade to *J. ailantifolia*, *J. mandshurica*, and *J. cathayensis* within section *Cardiocaryon* (Stanford et al. 2000; Aradhya et al. 2007); thus, *J. hopeiensis* can probably hybridize with any member of *Juglans* sect. *Cardiocaryon* that is introduced to its range, and with other *Juglans* species as well, as intersectional hybrids in *Juglans* are common (McGranahan and Leslie, 1991). Interestingly, Chen and Gilbert (2006) includes only three *Juglans* species (*J. regia*, *J. sigillata*, and *J. mandshurica*).

In this study, we assessed the genetic diversity, genetic structure, and demographic history of *J. hopeiensis*. The samples were evaluated at two mitochondrial sequences, two intergenic spacers of chloroplast DNA (cpDNA), the internal transcribed spacer region ITS1–ITS4, two polymorphic nuclear DNA sequences (*15R-8* and *Jr5680*), and 11 polymorphic nuclear microsatellites. Our specific aims were to (1) characterize the mitotypes, chlorotypes, and nuclear genetic variability of *J. hopeiensis* and (2), based on genetic and population genetic data, identify appropriate conservation strategies for *J. hopeiensis*.

Materials and methods

Sample collections and DNA extraction

Samples of leaves from *J. hopeiensis* (n = 48), *J. regia* (n = 30), and *J. mandshurica* (n = 30) were collected from 17 populations in China from 2013 to 2014 (Table 1). All sampled trees were healthy, mature specimens that appeared

Table 1 Sample information and haplotype summary for *Juglans hopeiensis*, *J. mandshurica*, and *J. regia* used in this study

Collection site	Pop ID	Species	Sample size	Longitude (E)	Latitude (N)	*3-9+ nad5*	*trn*S-G+*trn*L-F	*15R-8*	*Jr5680*	ITS
Fupingxian, Hebei	FP	*J. regia*	6	114.096	38.92	M4 (6)	H4 (6)	H9 (12)	H1 (12)	H14 (12)
Luliang, Shanxi	LL	*J. regia*	6	111.19	37.56	M4 (6)	H4 (6)	H9 (12)	H1 (12)	H14 (12)
Changzhi, Shanxi	CZ	*J. regia*	6	113.22	36.14	M4 (6)	H4 (6)	H9 (12)	H1 (12)	H14 (12)
Linfen, Shanxi	LF	*J. regia*	6	110.49	35.55	M4 (6)	H4 (6)	H9 (12)	H1 (12)	H14 (12)
Sanmenxia, Henan	SM	*J. regia*	6	111.03	34.80	M4 (6)	H4 (6)	H9 (12)	H1 (12)	H14 (12)
Laishui, Hebei	LS	*J. hopeiensis*	15	115.56	39.56	M1 (9), M2 (3), M3 (3)	H1 (8), H2 (7)	H1 (22), H2 (6), H3 (2),	H1 (14), H2 (1), H3 (1), H4 (1), H5 (2), H6 (11)	H1 (6), H2 (9), H3 (4), H4, (2) H5 (4), H6 (4), H7 (1)
Xinglong, Hebei	XL	*J. hopeiensis*	9	117.53	40.26	M1 (8), M3 (1)	H1 (9)	H1 (12), H3 (2), H4 (4)	H1 (8), H3 (1), H7 (4), H8 (5)	H2 (3), H3 (3), H4 (4), H5 (3), H8 (2), H9 (1), H10 (1), H11 (1)
Xiakou, Beijing	XK	*J. hopeiensis*	3	116.24	40.36	M1 (3)	H1 (3)	H1 (6)	H1 (1), H3 (1), H10 (1), H12 (1), H13 (1), H14 (1)	H2 (1), H11 (1), H15 (2), H17 (2)
Kuancheng, Hebei	KC	*J. hopeiensis*	3	118.38	40.60	M1 (1), M3 (1), M4 (1)	H1 (2), H3 (1)	H1 (4), H5 (1), H6 (1)	H1 (2), H9 (1), H10 (1), H11 (2)	H12 (1), H13 (1), H14 (2), H15 (1), H16 (1)
Zhuwo, Beijing	ZW	*J. hopeiensis*	4	115.76	40.09	M1 (1), M2 (2), M5 (1)	H1 (4)	H3 (4), H7 (4)	H12 (8)	H15 (5), H18 (3)
Panshan, Tianjin	PS	*J. hopeiensis*	2	117.28	40.10	M1 (2)	H2 (2)	H1 (2), H3 (2)	H12 (4)	H15 (2), H19 (2)
Xiaolongmen, Beijing	XM	*J. hopeiensis*	12	115.45	39.98	M1 (9), M5 (3)	H1 (12)	H1 (6), H2 (2), H3 (2), H4 (4), H7 (8), H8 (2)	H12 (24)	H15 (24)
Changbaishan, Jilin	AS	*J. mandshurica*	6	128.18	42.19	M5 (6)	H1 (6)	H2 (12)	H12 (12)	H20 (12)
Xiakou, Beijing	XC	*J. mandshurica*	6	116.25	40.35	M5 (6)	H1 (6)	H2 (12)	H12 (12)	H20 (12)
Miyun, Beijing	MY	*J. mandshurica*	6	116.96	40.44	M5 (6)	H1 (6)	H2 (12)	H12 (12)	H20 (12)
Yixian, Liaoning	YX	*J. mandshurica*	6	121.57	41.55	M5 (6)	H1 (6)	H2 (12)	H12 (12)	H20 (12)
Jianchang, Liaoning	JC	*J. mandshurica*	6	119.80	40.76	M5 (6)	H1 (6)	H2 (12)	H12 (12)	H20 (12)
Total			108							

Pop ID indicates that the name of populations, M indicates mitotypes, and H indicates haplotypes. Parentheses enclose number of individuals with the indicated haplotype or mitotype. *3-9* and *nad5* indicate the names of mitochondrial DNA fragment marker. *3-9+ nad5* indicates that we combined the two loci in our analysis. *trn*S-G and *trn*L-F indicate the names of chloroplast DNA markers. *trn*S-G+*trn*L-F indicates that we combined the two loci in our analysis. *15R-8*, *Jr5680*, and ITS indicate the names of nucleotide DNA fragment markers (for details, see Table S1).

to be autochthonous. We identified and collected samples of these three *Juglans* species based on their leaf, flower, or fruit morphology (Lu et al. 1999). The seven populations of *J. hopeiensis* that we collected covered the entire geographic distribution of the species (Table 1). Although the number of samples from some sites was small, the sites were exhaustively sampled. DNA was extracted following the methods of Doyle and Doyle (1987) and Zhao and Woeste (2011), and stored at −20 °C.

Mitochondrial, chloroplast, and nuclear DNA sequence analysis

The mitochondrial DNA sequences *3-9* (a likely *trn*H) (Zhao and Woeste 2011) and *nad5* (Dumolin-Lapegue et al. 1997), two chloroplast DNA segments (*trn*L-F, Zhao and Woeste, 2011; *trn*S-G, Zhang et al. 2005), and three nuclear loci (ITS, Zhao and Woeste 2011; *15R-8* (=GU552442, in *J. regia* = gb|LIHL01055671.1|) Zhao and Woeste 2011; *Jr5680*, a phenylalanine-ammonia lyase gene from *J. regia*, Dang et al. 2016) were polymorphic in all *Juglans* species in which they have been examined, so we analyzed these segments from each genome using a standard set of primers (Table S1).

The sequence data were edited and aligned using Bioedit v7.0.9 (Hall 1999). DnaSP v5.0 (Librado and Rozas 2009) was used to calculate the number of segregating sites, number of haplotype/mitotypes, mean number of pairwise differences (K), nucleotide diversity (pi), gene diversity within populations (h_S), total gene diversity (h_T), and parsimony informative sites. Each haplotype/mitotype was divided into sample site contributions and displayed as pie diagrams. We resolved the phased nuclear DNA sequences by applying the PHASE algorithms (Stephens and Donnelly 2003) in the software package DNASP v5.0 (Librado and Rozas 2009). Phylogenetic relationships between haplotypes were determined by constructing median-joining networks with Network v4.2.0.1 (Bandelt et al. 1999). Tajima's D (Tajima 1989) was used to examine the selective neutrality with significance tests based on 1000 permutations using Arlequin v3.11 (Excoffier 2007). The neutrality test statistics Fu and Li's D (Fu and Li 1993) and Fu's Fs were used to detect departures from mutation-drift equilibrium. The population genetic differentiation (F_{ST}) was determined with an analysis of molecular variance (AMOVA, Dupanloup et al. 2002); deviations from null distributions were tested with non-parametric permutation procedures (N = 99,999). To test whether the populations had undergone recent population growth, we plotted the mismatch distribution using the observed number of differences between pairs of haplotypes (Mousset et al. 2004).

Microsatellite data analysis

A total of 11 microsatellites (expressed sequence tag-simple sequence repeat (EST-SSR)) were designed using sequence data from *J. hopeiensis*, *J. mandshurica*, *J. cathayensis*, and *J. regia* (Hu et al. 2015, 2016; Dang et al. 2015, 2016; Table S2). All the EST-SSR-containing unigene sequences were BLAST (Basic Local Alignment Search Tool) searched in the NCBI database to identify their genic context (Table S3). We performed the PCR amplification of primer pairs on a Veriti 96-Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA) using 25-μL reactions. The upper primers were labeled with fluorescent dye, 6-FAM, HEX, TAMRA, and ROX (Sangon, Shanghai, China). The PCR products were visualized by an ABI 3730 sequencer (Applied Biosystems, USA). The allele size were assessed using GeneMapper v3.7 (Applied Biosystems, USA).

Genetic diversity per locus and population were evaluated through the following descriptive summary statistics: number of alleles (N_A), observed (H_O) and expected (H_E) heterozygosity, and inbreeding coefficient (F_{IS}) using the program GenAlEx 6.5 (Peakall and Smouse 2012). Allelic richness for a sample size of four (*J. hopeiensis*) was estimated with HP-Rare (Kalinowski 2005). GENEPOP v1.2 (Raymond and Rousset 1995) was used to test the Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) for all loci. The program CERVUS v3.0 (Kalinowski et al. 2007) was used to calculate the polymorphic information content (PIC). MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) was used to detect null alleles.

Genetic differentiation of populations (F_{ST}) was tested using the program GENEPOP v1.2 (Raymond and Rousset 1995). The significance of F_{ST} was determined by permutation tests (10,000) using Arlequin v3.5 (Excoffier and Lischer 2010). STRUCTURE was run using 100,000 burn-in MCMC iterations, with a run length of 1,000,000 iterations, and ten replicates per run for K = 2 to 9 clusters with admixture model (Pritchard et al. 2000). The software STRUCTURE HARVESTER was used to calculated the optimal value of K (Earl, 2012) using the delta K criterion (Evanno et al. 2005). The inferred clusters were drawn as colored boxplots using program DISTRUCT (Rosenberg 2004). STRUCTURE was run using two datasets: all trees sampled from all species (Table 1) and a second dataset containing only *J. hopeiensis* samples showing <20% admixture with *J. regia* or *J. mandshurica* in the analysis using all species. The overall genetic variation within and among different trees was explored by principal coordinate analysis (*PCoA*) using GenAlEx 6.5 (Peakall and Smouse 2012). The IBD software (Bohonak 2002) was used to analyze the Mantel test of geographic distance and genetic distance based on the IBWS (Isolation by distance web service) method (Jensen et al. 2005). The software Bottleneck v 1.2.02 was used to detect

demographic bottlenecks in *J. hopeiensis* populations (Piry et al. 1999) by a possible significant heterozygosity excess.

Results

Analysis of nucleotide diversity

The samples contained three variable sites within the two mitochondrial loci, constituting five mitotypes (Table 1; Figs. 1a and S1). All five mitotypes were observed in *J. hopeiensis*, while *J. regia* had mitotype M4 only, and *J. mandshurica* had mitotype M5 only (Fig. 1a). We believe that our samples probably represent the entire variability within *J. hopeiensis*, but other haplotypes have been documented for *J. regia* and *J. mandshurica* (Hu et al. 2015). Although five mitotypes were found in *J. hopeiensis*, only mitotype M1 was common, and the other four (M2, M3, M4, and M5) were found in two or fewer individuals of this species (Table 1; Fig. S1).

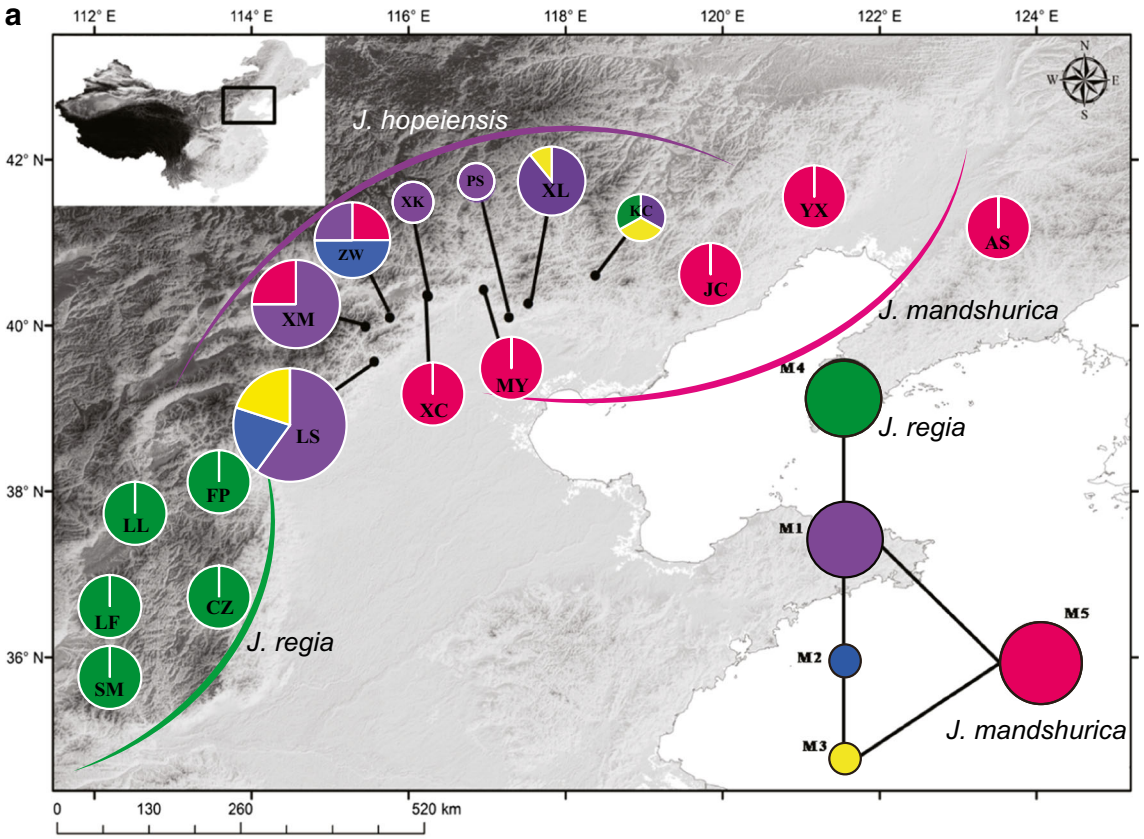
We obtained sequences for two cpDNA segments (aligned, edited length of 1558 bp) from 108 individuals in 17 populations. The total sequence contained 27 variable sites (Figs. 1b and S1) that constituted four haplotypes. Three haplotypes (H1, H2, and H3) were found in *J. hopeiensis* populations, while *J. regia* contained haplotype H4 only, and *J. mandshurica* contained haplotype H1 only (Fig. 1b). Only H1 and H4 were common' H2 and H3 occurred in five or fewer individuals (Figs. 1 and S1). The *trn*L-F DNA region presented the higher nucleotide diversity (0.0012), in spite of having the smaller number of nucleotides (1028 bp) and haplotypes (only two).

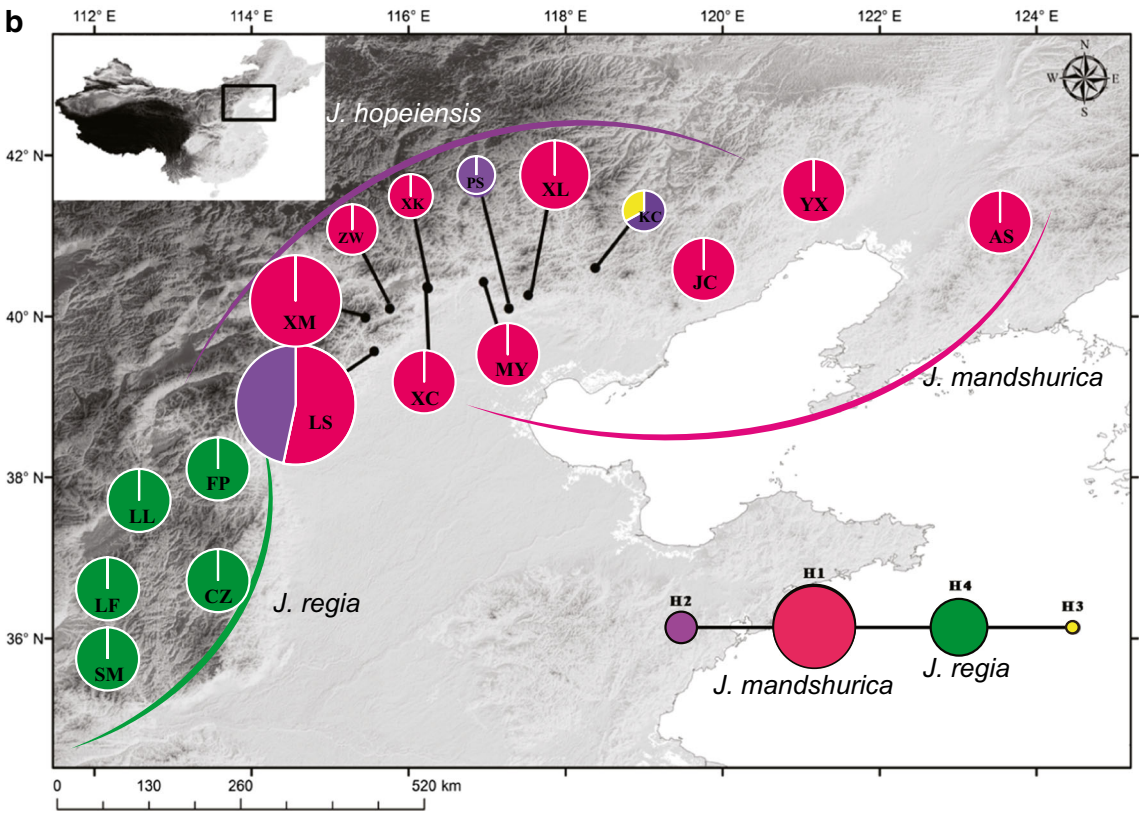
A total of 45 nuclear haplotypes were identified among the 108 individuals in three *Juglans* species based on sequence variation at the three nuclear segments that we analyzed (aligned, trimmed length of 1943 bp) (Fig. S1). A total of 69 variable sites were parsimony informative (Table 2). *J. regia* was monomorphic at *15R-8*, *Jr5680*, and ITS (haplotypes H9, H1, and H14, respectively), as was *J. mandshurica* (H2, H12, and H20) (Fig. 2). The *15R-8* DNA region (420 bp) presented the highest nucleotide diversity (0.0222), in spite of having the fewest haplotypes (nine) (Fig. 2a). The region, *Jr5680* (796 bp), contained 14 haplotypes but presented the lowest nucleotide diversity (0.0081) (Table 2; Fig. 2b). The ITS region contained 20 haplotypes, and its nucleotide diversity was 0.0161 (Table 2; Fig. 2c). *J. hopeiensis* presented 39 private haplotypes, while the remaining two haplotypes H2 (*15R-8*) and H12 (*Jr5680*) were found both in *J. hopeiensis* and *J. mandshurica* (Fig. 2a, b) and two haplotypes H9 (*15R-8*) and H1 (*Jr5680*) were found in both *J. hopeiensis* and *J. regia* (Fig. 2a, b). The haplotype H20 (ITS) was only found in *J. mandshurica*, while the haplotype H14 (ITS) was found in *J. regia* and *J. hopeiensis* (Fig. 2c).

Geographic structure, neutrality tests, and mismatch analysis

The geographic distribution of mitotypes revealed a phylogeographic signal that was somewhat stronger than the signal for cpDNA and nuclear ribosomal DNAs (nrDNAs), although there was a clear differentiation at the species level at the nuclear loci as well (Figs. 1 and 2). Haplotype H9 of *15R-8*, haplotype H1 of *Jr5680*, and haplotype H14 of ITS were found mostly not only in *J. regia* but also (2, 25, and 2 times, respectively) in 48 samples of *J. hopeiensis* and not at all in *J. mandshurica*. Conversely, only haplotype H2 of *15R-8* and H12 of *Jr5680* were found in *J. hopeiensis*, and H20 of ITS was found in *J. mandshurica*, and these haplotypes were not found in samples of *J. hopeiensis* or *J. regia* (Table 1; Fig. 2). For *J. hopeiensis*, the most common and the most widely distributed ancestral haplotype was H1 at locus *15R-8* (Fig. 2a). Haplotypes H2 to H11 and H14 of locus *Jr5680* were only distributed in *J. hopeiensis* (Fig. 2b). For ITS, haplotypes H14 and H20 were the only haplotypes in *J. regia* and *J. mandshurica*, respectively (Fig. 2c). Haplotype H15 was the ancestral haplotype for *J. hopeiensis* based on network analysis that included the other two *Juglans* species as outgroups. Network analysis of all three nuclear genes placed *J. hopeiensis* haplotypes as intermediate to *J. regia* and *J. mandshurica*. Maximum parsimony (MP) trees joined Ma walnut with *J. regia* as one group based on mitochondrial DNA (mtDNA) (bootstrap support <50), while Ma walnut was joined to *J. mandshurica* using cpDNA and nrDNA with strong bootstrap support (Fig. S2). The three walnut species (*J. hopeiensis*, *J. mandshurica*, and *J. regia*) shared mitotypes and chlorotypes, although no single haplotype was common to all three species and *J. hopeiensis* contained haplotypes not found in either *J. mandshurica* or Persian walnut (*J. regia*), and no mitotype or chlorotype was common to *J. mandshurica* and *J. regia* (Table 1). A similar pattern was also observed for the nuclear sequences (ITS, *15R-8*, and *Jr5680*).

Tests of neutrality for all nuclear, mitochondrial, and chloroplast sequences of *J. hopeiensis* were not conclusive. They revealed that Tajima's *D* was positive in all cases, Fu and Li's values were also positive, but none of the statistical tests were significant (Table 2). The mismatch distribution for the mitochondrial sequences had a single peak, which indicated that *J. hopeiensis* population size has expanded, experienced admixture with neighboring demes, or undergone a bottleneck (Fig. S3a). The mismatch distribution analyses for the cpDNA and nrDNA of *J. hopeiensis*, however, showed a double peak and multiple peaks, respectively, which indicated no recent expansion or contraction (Fig. S3b, c). Analyses using the software bottleneck showed that populations LS, XL, XK, and KC may have experienced a bottleneck (Table S4).


a
112° E
114° E
116° E
118° E
120° E
122° E
124° E
42° N
40° N
38° N
36° N
J. hopeiensis
J. mandshurica
J. regia
J. regia
J. mandshurica
PS
XL
XK
KC
YX
AS
ZW
JC
XM
MY
XC
LS
FP
LL
CZ
LF
SM
M4
M1
M5
M2
M3
0
130
260
520 km


b
112° E
114° E
116° E
118° E
120° E
122° E
124° E
42° N
40° N
38° N
36° N
J. hopeiensis
J. mandshurica
J. regia
J. regia
J. mandshurica
PS
XL
XK
KC
YX
AS
ZW
JC
XM
MY
XC
LS
FP
LL
CZ
LF
SM
H1
H2
H4
H3
0
130
260
520 km

◀ **Fig. 1** Geographical distribution of haplotypes and haplotype network for 17 populations of *J. hopeiensis*, *J. mandshurica*, and *J. regia* based on two mtDNA (*3-9* and *nad5*) sequences (**a**) and two cpDNA(*trn*S-G and *trn*L-F) sequences (**b**). *Colors* were assigned to each mitotype/haplotype according to the legend at the *right side* of the figure. Circumference of pie charts indicates the number of samples in each population. Population *symbols* are identified in Table 1. The samples of *J. hopeiensis* were collected from the species' entire range. Samples of Persian walnut (*J. regia*) were collected from region of Henan, Shanxi, and Hebei province in China. Samples of *J. mandshurica* were collected from Northeast China near Beijing. Phylogenetic relationships between mtDNA mitotypes or cpDNA haplotypes and evolutionary relationships within each species were determined by constructing median-joining networks (*inset*, *lower right*). The small map (*upper left*) shows the location of the studied area

Analysis of genetic diversity using microsatellites

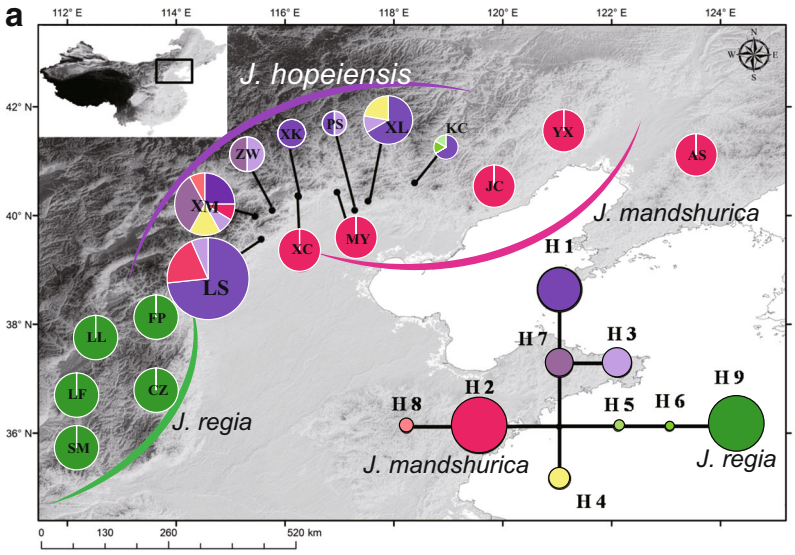
Analysis of seven populations of *J. hopeiensis* based on 11 polymorphic microsatellite loci showed that alleles per locus (N_A) ranged from 6 to 18 with a mean of 9.5. The observed heterozygosity (H_O) ranged from 0.15 to 0.62 (mean 0.38) and expected heterozygosity from 0.51 to 0.87 (mean 0.67). The polymorphic information content (PIC) ranged from 0.662 to 0.907 (mean 0.77). Shannon index (*I*) ranged from 1.35 to 2.22, with a mean of 1.70 (Table S2). MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) did not reveal any null alleles. When the same analyses were applied to all 17 populations (all three *Juglans* species), alleles per locus (N_A) ranged from 1.727 to 6.364 with a mean of 3.032. The observed heterozygosity (H_O) and expected heterozygosity (H_E) varied from 0.136 to 0.636 (mean = 0.386) and from 0.227 to 0.759 (mean = 0.463), respectively. Shannon's index (*I*) ranged from 0.371 to 1.597, with a mean of 0.824 (Table S5). Allelic richness (AR) ranged from 1.490 to 2.920, with a mean of 2.126. Allelic richness ranged from 1.490 to 1.820 for *J. regia*, allelic richness ranged from 1.910 to 2.920 for *J. hopeiensis*, and allelic richness ranged from 1.740 to 2.710 for *J. mandshurica*, respectively (Table S5). The population LS had ten private alleles, while XM, XL, KC, and ZW had three, one, one, and one private alleles, respectively (Tables S6 and S7). A large number of private alleles (ten) were identified across eight loci at the LS sample site. Private alleles at site XM were at loci JM61666 and JM5969 (three total alleles, frequencies were 0.083, 0.042, and 0.083, respectively); at site ZW, a private allele was at locus JR6439 (frequency was 0.125); at site KC, a private allele was at JM5969 (frequency was 0.167); and at XL, a private allele was at locus JR4616 (allele frequency was 0.056) (Table S6). Four populations showed significant departure from the Hardy-Weinberg equilibrium (HWE) for at least some loci, LS (eight loci, JR4616, JM61666, JM28820, JM78331, JC8125, JH89978, and JM77909), XM (seven loci, JR6439, JR3773, JM61666, JM28820, JM5969, JH84548, and JH89978), XL (five loci, JR3773,

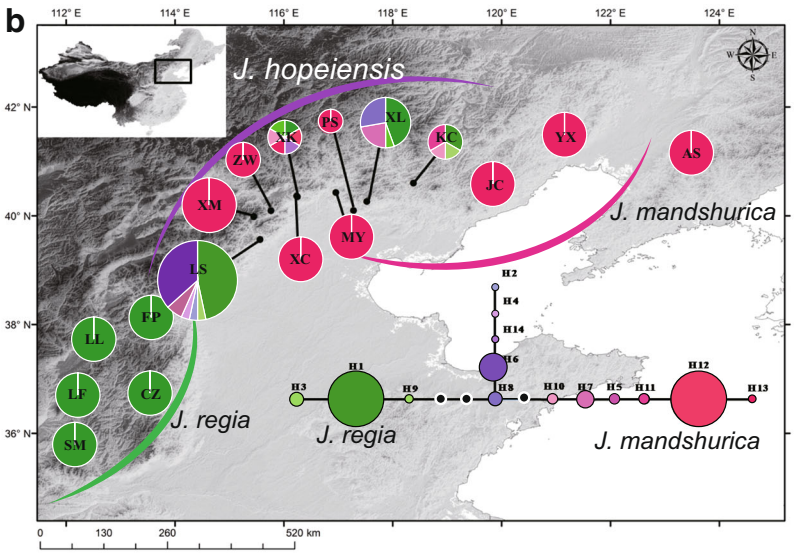
Table 2 Characterization and diversity of two mtDNA (*nad5* and *3-9*), two cpDNA (*trn*L-F and *trn*S-G), and three nrDNA (ITS, *15R-8*, and *Jr5680*) DNA sequences for 17 populations of *J. hopeiensis*, *J. mandshurica*, and *J. regia*

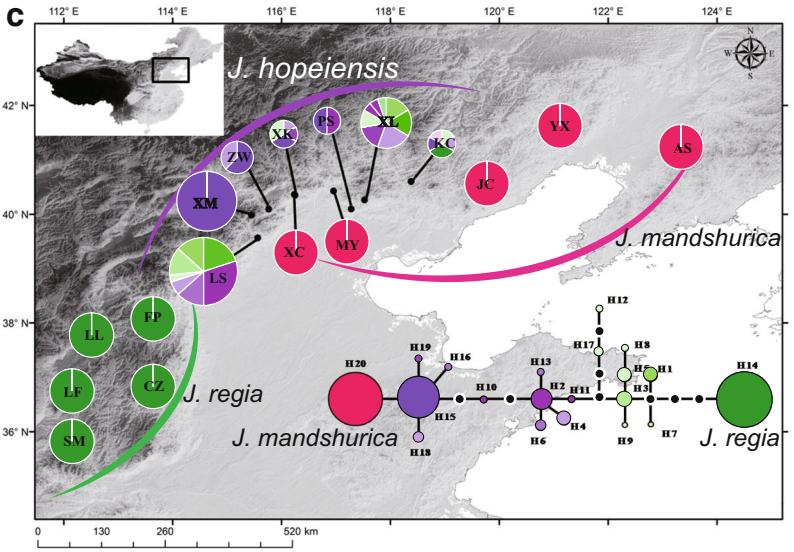
Type	Locus/ primer	Fragment length (bp)	Parsimony informative sites	Number of haplotypes	Number of polymorphic sites	Haplotype diversity (Hd)	Nucleotide diversity	Tajima's *D*	Fu and Li's *D*	Fu and Li's *F*
mtDNA	*3-9*	205	2	3	2	0.677	0.0044	1.191	0.873	1.132
	nad5	629	1	2	1	0.127	0.0002	$P > 0.10$	$P > 0.10$	$P > 0.10$
	All	834	3	5	3	0.718	0.0012			
cpDNA	*trn*S-*trn*G	530	3	4	3	0.532	0.0020	1.908	1.155	1.644
	*trn*L-*trn*F	1028	3	2	24	0.424	0.0012	$0.10 > P > 0.05$	$P > 0.10$	$0.10 > P > 0.05$
	All	1558	6	4	27	0.522	0.0015			
nrDNA	*15R-8*	420	26	9	26	0.782	0.0222	3.631	2.07	3.318
	Jr5680	796	13	14	14	0.677	0.0081	$P < 0.001$	$P < 0.02$	$P < 0.02$
	ITS1-ITS4	708	28	20	29	0.853	0.0161			
	All	1924	69	45	69	0.899	0.0140			

For details of primers, see Tables 1 and S1

mtDNA mitochondrial DNA, *cpDNA* chloroplast DNA fragment, *nrDNA* nuclear DNA


a
J. hopeiensis
J. mandshurica
J. regia
XK
PS
XL
KC
YX
AS
ZW
JC
XM
MY
XC
LS
FP
LL
LF
GZ
SM
H 1
H 7
H 3
H 9
H 2
H 8
H 5
H 6
H 4
112° E
114° E
116° E
118° E
120° E
122° E
124° E
42° N
40° N
38° N
36° N
0
130
260
520 km


b
J. hopeiensis
J. mandshurica
J. regia
XK
PS
XL
KC
YX
AS
ZW
JC
XM
MY
XC
LS
FP
LL
LF
GZ
SM
H2
H4
H14
H6
H1
H3
H9
H8
H10
H7
H5
H11
H12
H13
112° E
114° E
116° E
118° E
120° E
122° E
124° E
42° N
40° N
38° N
36° N
0
130
260
520 km


c
J. hopeiensis
J. mandshurica
J. regia
XK
PS
XL
KC
YX
AS
ZW
JC
XM
MY
XC
LS
FP
LL
LF
GZ
SM
H12
H17
H8
H19
H16
H20
H13
H10
H2
H11
H1
H3
H14
H15
H4
H6
H9
H7
H18
112° E
114° E
116° E
118° E
120° E
122° E
124° E
42° N
40° N
38° N
36° N
0
130
260
520 km

◀ **Fig. 2** Geographical distribution of haplotypes and network of haplotype for 17 populations of *J. hopeiensis*, *J. mandshurica*, and *J. regia* based on three nrDNA (*15R-8*, *Jr5680*, and ITS) sequences. Patterns were assigned for each haplotype according to the network at the *right side* of the figure. **a** Results for nrDNA sequence *15R-8*. **b** Results for nrDNA sequence *Jr5680*. **c** Results for nrDNA sequence ITS

JM61666, JM68820, JC8125, and JM77909), and ZW (two loci, JM5969 and JH89978). We observed no isolation by distance among *J. hopeiensis* populations. Applying the Mantel test for matrix correlation did not reveal any correlation between genetic distance and geographic distance ($r = 0.007$, $P = 0.530$, Fig. S5).

Analysis of genetic structure

When all samples (from all three species) were included in the STRUCTURE analysis (which was based on EST-SSR data only), the highest mean posterior probability value was detected for $K = 3$ (Fig. 3). For $K = 3$, the proportion of membership in each population for each individual strongly reflected

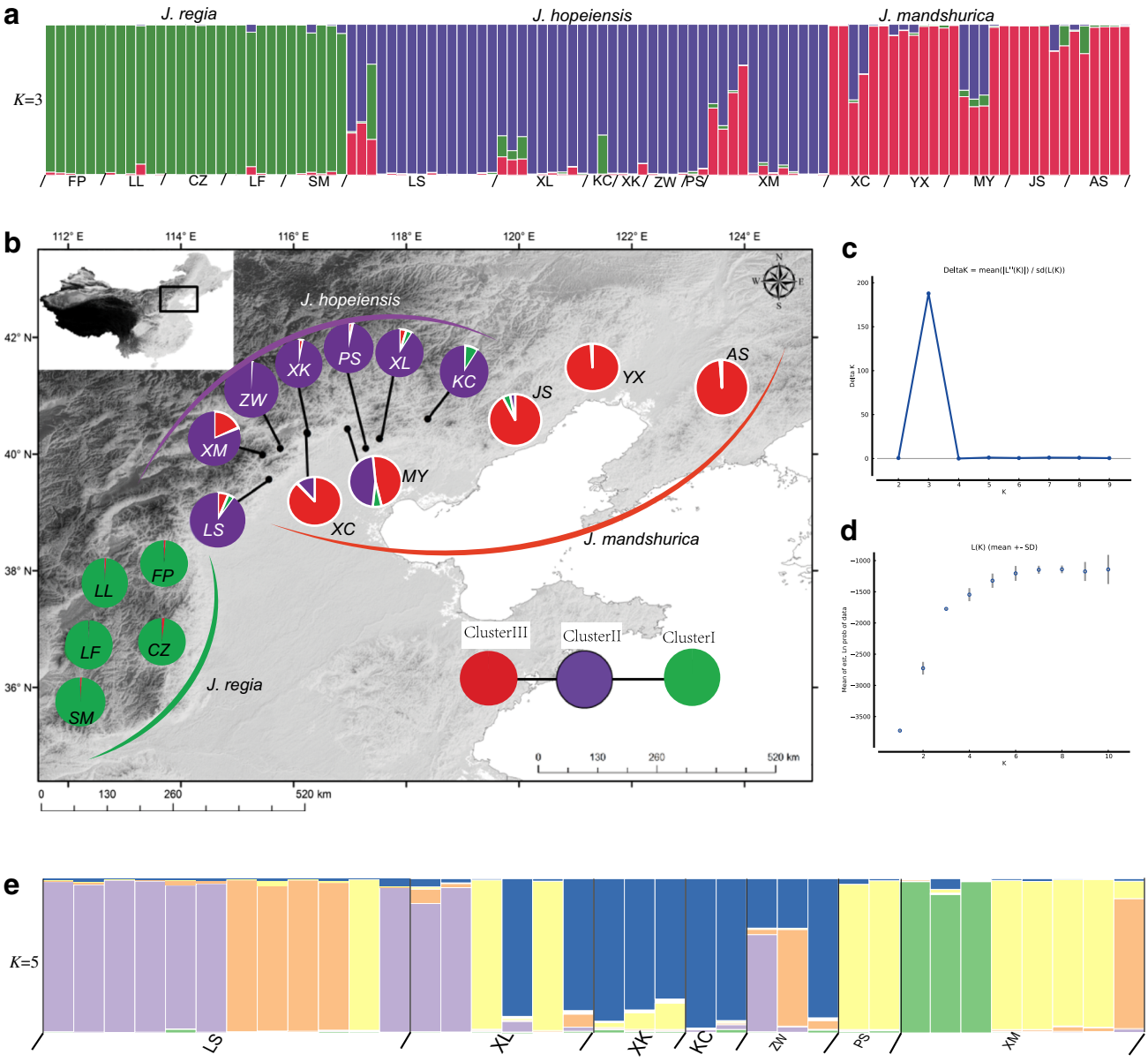


Fig. 3 Probability of membership in each of the three species (Q-plot) based on Bayesian model-based clustering of 108 individuals of *J. hopeiensis*, *J. mandshurica*, and *J. regia* using 11 EST-SSRs. **a** Estimated genetic clustering ($K = 3$) obtained with the STRUCTURE program for 108 individuals. Individuals are separated according to the population, and the *black vertical line* in the bar chart is population identifier. **b** Location of each sampled population and proportion of admixture of *J. hopeiensis*, *J. regia*, and *J. mandshurica* in each population are indicated with pie charts. **c** Distribution of delta *K* for $K = 2$ to 9 to determine the true number of populations (*K*) as described in Evanno et al. (2005). **d** Mean log likelihood of the data at varying estimates of *K*. **e** Structure analysis of *J. hopeiensis* sampled from seven locations showing moderate differentiation ($K=5$) and little admixture

species membership, as expected. At $K = 3$, some admixture between *J. hopeiensis* and *J. mandshurica* was indicated. This pattern was also detected in the core dataset of 11 EST-SSR loci used in the *PCoA* analysis, as shown in Fig. S4. In the maximum parsimony tree based on mitochondrial sequence, *J. hopeiensis* occupied a separate branch from *J. mandshurica* and *J. regia* (Fig. S2a). At the nuclear level, however, an overlap between *J. mandshurica* and *J. hopeiensis* was obvious, as was their separation from *J. regia* (except for a small number of haplotypes that probably reflect hybrid ancestry). STRUCTURE, a Bayesian clustering software program, separated the samples of *J. hopeiensis* into five populations based on EST-SSRs (Fig. 3e).

Analysis of molecular variance

Microsatellites revealed a clear genetic distinction among the seven *J. hopeiensis* populations studied ($F_{ST} = 0.097$, $P < 0.001$). These results were corroborated by analysis of molecular variance (AMOVA) (Table 3). AMOVA analysis of the data using the mtDNA, cpDNA, and *15R-8*, ITS, and *Jr5680* sequences of *J. hopeiensis* populations showed that the coefficient of genetic differentiation at these loci was (F_{ST}) was −0.030 ($P < 0.05$), 0.191 ($P < 0.01$), 0.156 ($P < 0.01$), 0.373 ($P < 0.01$), and 0.625 ($P < 0.01$), respectively (Table 3).

Discussion

Analysis of genetic structure, the presence of private alleles and haplotypes, and the MP tree based on mitochondrial marker data (Fig. S2a) all indicated that *J. hopeiensis* is sister to *J. mandshurica*, but is a distinct evolutionary lineage. In the Bayesian STRUCTURE analysis based on EST-SSRs, samples from all seven locations (representing the entire range of the species) clustered into a single group distinct from *J. regia* and *J. mandshurica* (Fig. 3). But the STRUCTURE analysis also showed that *J. hopeiensis* may be undergoing genetic swamping by admixture with both *J. regia* and *J. mandshurica*. Genetic swamping and introgression by related taxa can threaten species' genetic and evolutionary integrity (Holderegger et al. 2005), and it represents an important hazard to plant conservation (Ellstrand 1992). How hybridization affects *J. hopeiensis* specifically (in terms of fitness) has not been determined but is an important research question. Certainly human dispersal and selection have played

Table 3 AMOVA results for seven populations of *J. hopeiensis* using multiple loci

Type of genetic markers	Source	*df*	SS	Est. Var.	*P* (%)
EST-SSR	Among populations	6	56.03	0.43 Va	9.69
	Within populations	89	353.12	3.98 Vb	90.31
	Total	95	409.16	4.39	100
	$F_{ST} = 0.097$				
mtDNA (*3-9+nad5*)	Among populations	6	1.67	−0.01 Va	−3.04
	Within populations	17	5.25	0.31 Vb	103.04
	Total	23	6.92	0.30	100
	$F_{ST} = -0.030$				
cpDNA (*trn*L-F+*trn*S-G)	Among populations	6	10.35	0.21 Va	19.06
	Within populations	21	18.83	0.90 Vb	80.94
	Total	27	29.18	1.11	100
	$F_{ST} = 0.191$				
nrDNA (*15R-8*)	Among populations	6	36.01	0.33 Va	15.65
	Within populations	89	159.98	1.80 Vb	84.35
	Total	95	195.99	2.13	100
	$F_{ST} = 0.156$				
nrDNA (*ITS*)	Among populations	6	164.95	1.92 Va	37.30
	Within populations	89	287.79	3.23 Vb	62.70
	Total	95	452.74	5.16	100
	$F_{ST} = 0.373$				
nrDNA (*Jr5680*)	Among populations	6	176.42	2.23 Va	62.47
	Within populations	89	118.99	1.34 Vb	37.53
	Total	95	295.41	3.56	100
	$F_{ST} = 0.625$				

df degree freedom, *SS* deviation square, *MS* mean square, *Est. Var.* variance components, *P* percentages of molecular variance

a role in maintaining *J. hopeiensis* on the landscape, but they have also encouraged its hybridization, which likely increases its genetic diversity but may not increase its fitness (Vilà et al. 2000; Crystal et al. 2016). Outbreeding depression and lack of adaptation are possible consequences of interspecific hybridization (Goto et al. 2011; Bleeker et al. 2007).

J. hopeiensis populations not only contained unique haplotypes, but also shared haplotypes at nuclear loci, mitotypes, and chlorotypes of *J. regia* and *J. mandshurica* (Figs. 1, 2, and S2). *J. hopeiensis* also contained microsatellite alleles not found in the other two *Juglans* species (Table S6). It is possible that nuclear sequences, chlorotypes, and mitotypes that were unique to *J. hopeiensis* in our dataset are also found in *J. regia*, *J. mandshurica*, or even *J. cathayensis* or *J. sigillata* (Iron walnut) at locations that we did not sample. For example, chlorotype H3 of *J. hopeiensis*, which appeared to be derived from the only observed *J. regia* chlorotype (H4, Fig. 1b network), was found in only a single individual of population KC (Fig. 1b). This individual also contained mitotype M4 (found in all *J. regia*), showed considerable admixture in the nuclear genome with *J. regia* based on EST-SSR alleles (Fig. 3), and contained the nuclear haplotypes H28 and H29 that were least differentiated from *J. regia* based on maximum parsimony (Fig. S2). These data may indicate that chlorotype H3 was actually a *J. regia* chlorotype found in a cryptic *J. regia* × *J. hopeiensis* hybrid individual, which may explain why H3 is not connected to H2 (the only other *J. hopeiensis* chlorotype) in the network diagram of Fig. 1b and why the same individuals showed H14 (a *J. regia* haplotype at ITS, Fig. 2c) and H5 and H6 at nuclear locus *15R-8*, two haplotypes not found in the *J. regia* samples, but closely related to them (Fig. 2a). Similarly, H8 of nuclear locus *15R-8* (Fig. 2a) was identified in three individuals of population XM. Haplotype H8 was closest genetically to H2, the haplotype common to all *J. mandshurica* that we sampled and disconnected from other *J. hopeiensis* haplotypes. These three individuals appeared highly admixed at EST-SSR loci (Fig. 3a) and contained mitotype M5, which was common to all *J. mandshurica* that we sampled (Fig. 1a). It seems likely that samples XM-4, XM-5, and XM-6 are *J. mandshurica* × *J. hopeiensis* hybrids. A conclusion of hybridity for these individuals is uncertain, however, because other *J. hopeiensis* individuals with an M5 mitotype did not show strong admixture at EST-SSRs (Fig. 3a).

Our analyses showed that *J. hopeiensis* was closer genetically to *J. mandshurica* than *J. regia*, especially when nuclear sequence data were considered (Figs. 1, 2, and S2). This is not a surprising result, for *J. regia* cannot be considered native to the parts of China containing *J. hopeiensis*, but it has grown there as a cultivated species for many generations (Chen et al. 2014), and *J. hopeiensis* has long been considered a member of sect. *Cardiocaryon*, rather than sect. *Dioscaryon* of *Juglans*. The reduction in population size of the *J. hopeiensis* may be due to genetic swamping caused by pollen flow form other cultivated *Juglans* species in nearby farms (Millar et al. 2012) and other processes that mediate hybridization between closely related species (Hoban et al. 2012; Gómez et al. 2015; Shu et al. 2016). Some of the nuclear haplotypes of *J. hopeiensis* that clustered near to *J. mandshurica* based on maximum parsimony (Fig. S2) were identified in individuals that also reflected admixture at microsatellite loci. For example, H41 was found in the highly admixed samples XM-1 and XM-3 (Fig. 3), H42 was found in XM-4, and H37 was found in XM-2. Nevertheless, there was no perfect correspondence between the Bayesian analyses based on microsatellites (Fig. 3) and the maximum parsimony branch placement (Fig. S2). Some samples with relatively greater admixture in the STRUCTURE analysis did not show evidence of hybridization based on sequence data from any of the three nuclear loci, so their haplotypes were strongly differentiated from *J. regia* and *J. mandshurica* in the maximum parsimony tree (Fig. S2).

Both the nrDNA sequence data and EST-SSR data showed that there is probably a high level of gene flow between *J. hopeiensis* and *J. mandshurica*, presumably though wind pollination and possibly recently by human selection for thick, ornate shells. The spatially explicit maps (Figs. 1, 2, and 3) show that *J. hopeiensis* is fully sympatric with *J. mandshurica*, and the results from STRUCTURE show that gene flow between *J. hopeiensis* and *J. mandshurica* is more common and probably more biologically important than between *J. regia* and either of the other two species (Fig. 3). Only *J. hopeiensis* populations ZW and PS showed few clear signs of introgression with other *Juglans* species, but sample sizes from these populations were small, so their apparent genetic purity may be a sampling artifact. Some *J. hopeiensis* samples appeared to be complex hybrids involving *J. regia* and *J. mandshurica* or possibly other *Juglans* species.

The genetic features and genetic structure of *J. hopeiensis* have important implications for the conservation and management of this narrowly distributed and extremely rare species. Because the species is rare, all existing populations should be protected from exploitation and their natural regeneration encouraged, but particular features of some populations deserve to be highlighted here. Populations LS, XL, KC, ZW, and XM contained private alleles at nuclear loci (Table S7), and population PS was the only population in which the chlorotype of all sampled individuals was H2, the chlorotype unique to *J. hopeiensis* (Fig. 2b). Within *J. hopeiensis*, only chlorotype H3 was private to one population; H3 was observed in a single individual of population KC and is probably a *J. regia* chlorotype. Private alleles may represent important genetic resources for species survival. It is likely that there has already been allele loss from three of these populations (LS, XL, and KC), based on genetic bottleneck analysis (Table S4). A sudden and dramatic reduction in population size may have been the result of the use of *J. hopeiensis* for timber or for gunstocks

during wars of the last century (Pei et al. 2006). From 1949 to 1978, Ma walnut received almost no scientific attention because the value of its nuts was low, it grows relatively slowly, and its natural reproductive capacity is low (Pei et al. 2006). With the improvement of living standards in China, Ma walnut is increasingly sought as a curio or talisman, and the nuts are even exported (Wu et al. 1999; Pei et al. 2006). As a consequence, resources of this species are suffering from long-term predatory development, illegal logging, overgrazing, and habitat fragmentation, all of which threaten the species with extinction (Chen et al. 2014; Pei et al. 2006).

Compared to the (geographically restricted) samples of the other two species, *J. hopeiensis* was considerably more genetically diverse, especially populations XM, LS, and XL (populations with the largest sample sizes, Table S5; Fig. 2a), which may have been a reflection of the number of trees sampled from those sites. Populations LS, XL, and XM were the most genetically diverse (Table S5) based on effective numbers of alleles and unbiased expected heterozygosity. Population LS also contained a large number of private alleles, as previously mentioned. The accuracy of our within-group diversity parameter estimates (Table S5) is low, since we were constrained by small sample sizes. These estimates should be considered rough approximations. These estimates do support the general impression that the genetic diversity of *J. hopeiensis* is comparable to *J. regia* and *J. mandshurica*, two species not considered at risk. Further sampling will be needed to determine the extent to which higher levels of diversity in these populations, as compared to the others, were biased by sample sizes and/or by admixture with other *Juglans* species. For example, some individuals with mitotype M5 of populations XM might be *J. mandshurica* hybrids (Fig. 1a). A low level of genetic diversity is commonly expected for narrow endemic plants (Ellstrand and Elam 1993), but that is not always the case (Fernández-Mazuecos et al. 2014). Walnuts are large, long-lived trees with a heterodichogamous (highly outcrossing) mating system (Kimura et al. 2003). These life traits and this reproductive strategy typically maintain high levels of genetic diversity within populations. We observed that genetic diversity measures for *J. hopeiensis* populations were similar to or even higher than those for *J. regia* and *J. mandshurica* (Table S5), species not considered at risk. Our sample sizes from *J. regia* and *J. mandshurica* were small and geographically restricted, however, so our estimates of diversity for these species are likely biased.

The natural distribution of *J. hopeiensis* in China is limited to a series of separated populations. The seven locations that we sampled revealed relatively little interpopulation hybridization (admixture between *J. hopeiensis* populations), which probably means that they are not exchanging much pollen (Fig. 3e). The relatively high genetic differentiation among populations of *J. hopeiensis* (Table 3; Fig. 3e), relatively large numbers of private nuclear alleles present at relatively high frequencies (Table S6), large number of loci significantly out of HWE (Table S2), and many loci with high F_{IS} values (Table S2) may all indicate a lack of gene flow within and among demes. The level of population differentiation (F_{ST} = 0.097) of *J. hopeiensis* was slightly higher than that observed in many wild, wind-pollinated *Juglans* species with continuous distributions analyzed using SSRs (North American species), but comparable to Chinese *Juglans* species with wild populations that are often disjunct (*J. mandshurica*, F_{ST} = 0.054, Wang et al. 2016; *J. regia*, F_{ST} = 0.16, Pollegioni et al. 2015; *Juglans cathayenesis*, F_{ST} = 0.110, Bai et al. 2014; *J. sigillata*, F_{ST} = 0.111, Wang et al. 2015), whereas *Juglans* in eastern North America grow in forests that are less disjunct than those in Asia (*Juglans nigra*, F_{ST} = 0.017, Victory et al. 2006; *Juglans cinerea*, F_{ST} = 0.045, Hoban et al. 2010). Many factors, including population size, sample sizes, and biogeography, can influence estimates of F_{ST}, but data from EST-SSRs do not strongly indicate that *J. hopeiensis* populations are more shaped by isolation than any other Chinese *Juglans* species.

Small populations are especially susceptible to loss (Matthies et al. 2004). Therefore, we recommend that in situ conservation efforts should be implemented especially for the unprotected small populations PS, XK, and KC, to prevent their further deterioration. Efforts toward enrichment or enhancement of natural regeneration should be considered. Trees in all populations should be sampled to determine the likelihood that they are hybrids and whether they should be removed or repropagated to another location. The identification of any other natural populations of *J. hopeiensis* and their risk of gene admixture from other *Juglans* should also be a high priority. *J. hopeiensis* is sympatric with *J. mandshurica*, and *J. regia* has been introduced to most if not all of the *J. hopeiensis* range. The role of interspecific hybridization in the maintenance of genetic diversity and/or the loss of fitness in *J. hopeiensis* needs further investigation. If conspecific gene flow results in outbreeding depression, then populations should be managed to reduce gene flow (Ellstrand 1992). *J. hopeiensis* in living collections such as arboreta or botanical parks should be genotyped to verify its taxonomic status and as a potential source of pollen for controlled crosses. The risk of loss of *J. hopeiensis* by genetic swamping needs to be evaluated thoroughly. Isolation of nonadmixed *J. hopeiensis* should take place by removal of *J. mandshurica* and *J. regia* from within a buffer zone of several kilometers. Much more data are needed to craft a comprehensive conservation plan for *J. hopeiensis*; however, we believe that the data presented here represent an important starting point.

Acknowledgements The authors wish to thank Shuhai Yin and Yong Wang for sample collection. This work was supported by the National Natural Science Foundation of China (41471038; 31200500; J1210063), the Training Programs of Innovation for graduates (YZZ15062), and

Changjiang Scholars and Innovative Research Team in University (IRT1174). Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the US Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that also may be suitable.

Data archiving statement The haplotype sequences were deposited in GenBank under accession numbers KY660646–KY660676 and KY652952–KY652972.

References

Aradhya MK, Potter D, Gao F, Simon CJ (2007) Molecular phylogeny of *Juglans* (Juglandaceae): a biogeographic perspective. Tree Genet Genomes 3:363–378

Bai WN, Liao WJ, Zhang DY (2010) Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. New Phytol 188(3):892–901

Bai WN, Wang WT, Zhang DY (2014) Contrasts between the phylogeographic patterns of chloroplast and nuclear DNA highlight a role for pollen-mediated gene flow in preventing population divergence in an East Asian temperate tree. Mol Phylogenet Evol 81:37–48

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Bleeker W, Schmitz U, Ristow M (2007) Interspecific hybridisation between alien and native plant species in Germany and its consequences for native biodiversity. Biol Conserv 137:248–253

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. J Hered 93:153–154

Broadhurst LM, Lowe A, Coates DJ, Cunningham SA, McDonald M, Vesk PA, Yates C (2008) Seed supply for broadscale restoration: maximizing evolutionary potential. Evol Appl 1:587–597

Budd C, Zimmer E, Freeland JR (2015) Conservation genetics of *Magnolia acuminata*, an endangered species in Canada: can genetic diversity be maintained in fragmented, peripheral populations? Conserv Genet 16:1359–1373

Chen L, Ma Q, Chen Y, Wang B, Pei D (2014) Identification of major walnut cultivars grown in China based on nut phenotypes and SSR markers. Sci Hortic Amsterdam 168:240–248

Chen SL, Gilbert MG (2006) Flora of China. Science, Beijing and Missouri Botanical Garden Press, St Louis

Crystal PA, Lichti NI, Woeste KE, Jacobs DF (2016) Vegetative and adaptive traits predict different outcomes for restoration using hybrids. Front Plant Sci 7:1741

Dang M, Liu ZX, Chen X, Zhang T, Zhou HJ, Hu YH, Zhao P (2015) Identification, development, and application of 12 polymorphic EST-SSR markers for an endemic Chinese walnut (*Juglans cathayensis* L.) using next-generation sequencing technology. Biochem Syst Ecol 60:74–80

Dang M, Zhang T, Hu Y, Zhou H, Woeste KE, Zhao P (2016) De novo assembly and characterization of bud, leaf and flowers transcriptome from *Juglans regia* L. for the identification and characterization of new EST-SSRs. Forests 7:247

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Dumolin-Lapegue S, Pemonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. Mol Ecol 6:393–397

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571–2581

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Ellstrand NC (1992) Gene flow by pollen: implications for plant conservation genetics. Oikos 63:77–86

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. Annu Rev Ecol Syst 24:217–242

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L (2007) ARLEQUIN, version 3.11. Computational and Molecular Population Genetics Lab, CMPG. Zoological Institute University of Berne, Switzerland

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

Fernández-Mazuecos M, Jiménez-Mejías P, Rotlan-Puig X, Vargas P (2014) Narrow endemics to Mediterranean islands: moderate genetic diversity but narrow climatic niche of the ancient, critically endangered *Naufraga* (Apiaceae). Perspect Plant Ecol 16:190–202

Fjellstrom RG, Parfitt DE (1995) Phylogenetic analysis and evolution of the genus *Juglans* (Juglandaceae) as determined from nuclear genome RFLPs. Plant Syst Evol 197:19–32

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

Gómez JM, González-Megías A, Lorite J, Abdelaziz M, Perfectti F (2015) The silent extinction: climate change and the potential hybridization-mediated extinction of endemic high-mountain plants. Biodivers Conserv 24:1843–1857

Goto S, Iijima H, Ogawa H, Ohya K (2011) Outbreeding depression caused by intraspecific hybridization between local and nonlocal genotypes in *Abies sachalinensis*. Restor Ecol 19:243–250

Gunn BF, Aradhya M, Salick JM, Miller AJ, Yongping Y, Lin L, Xian H (2010) Genetic variation in walnuts (*Juglans regia* and J. sigillata; Juglandaceae): species distinctions, human impacts, and the conservation of agrobiodiversity in Yunnan, China. Am J Bot 97:660–671

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hao Y, Huang W, Wang K, Qi J, Xu J (2007) Analysis of Sect. *Juglans* germplasm in China using SSR marker. J Fruit Sci 5:012 **(In Chinese)**

Hao YB, Wu CL, Chen YH, Dong NG, Wang WX, Qi JX, Pei D (2013) Selection of a new *Juglans hopeiensis* cultivar—'Jingyi 1'. J Fruit Sci 30:718–719 **(In Chinese)**

He N, Ning D, Xu T, Ma T, Li Y (2015) Early growth performance of introduction and cultivation of *Juglans hopeiensis* Hu in Yunnan province. For Invent Plan 40:106–109 **(In Chinese)**

Hoban SM, Borkowski DS, Brosi SL, McCleary TIM, Thompson LM, McLachlan JS, Pereira M, Schlarbaum SE, Romero-Severson JEANNE (2010) Range-wide distribution of genetic diversity in the north American tree *Juglans cinerea*: a product of range shifts, not ecological marginality or recent population decline. Mol Ecol 19:4876–4891

Hoban SM, McCleary TS, Schlarbaum SE, Anagnostakis SL, Romero-Severson J (2012) Human-impacted landscapes facilitate hybridization between a native and an introduced tree. Evol Appl 5:720–731

Holderegger R, Angelone S,·Brodbeck S,·Csencsics D,·Gugerli F, ·Hoebee SE,·Finkeldey R (2005) Application of genetic markers to the discrimination of European black poplar (*Populus nigra*) from American black poplar (*P. deltoides*) and hybrid poplars (*P. x canadensis*) in Switzerland. Trees 19:742–747

Hu YH, Zhao P, Zhang Q, Wang Y, Gao XX, Zhang T, Zhou HJ, Dang M, Woeste KE (2015) *De novo* assembly and characterization of transcriptome using Illumina sequencing and development of twenty

five microsatellite markers for an endemic tree *Juglans hopeiensis* Hu in China. Biochem Syst Ecol 63:201–211

Hu Z, Zhang T, Gao XX, Wang Y, Zhang Q, Zhou HJ, Zhao GF, Wang ML, Woeste KE, Zhao P (2016) *De novo* assembly and characterization of the leaf, bud, and fruit transcriptome from the vulnerable tree *Juglans mandshurica* for the development of 20 new microsatellite markers using Illumina sequencing. Mol Gen Genomics 291:849–862

Hu YH, Woeste KE, Zhao P (2017) Completion of the chloroplast genomes of five Chinese *Juglans* and their contribution to chloroplast phylogeny. Front Plant Sci 7:1955

Jensen J, Bohonak AJ, Kelley SK (2005) Isolation by distance Web service. BMC Genet 6:13

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5: 187–189

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106

Kimura M, Seiwa K, Suyama Y, Ueno N (2003) Flowering system of heterodichogamous *Juglans ailanthifolia*. Plant Spec Biol 18:75–84

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Lu AM (1982) The geographical dispersal of Juglandaceae. Acta Phytotaxon Sin 20:257–274

Lu AM, Stone DE, Grauke LJ (1999) Juglandaceae. Flora China 4:277–285

Manning WE (1978) The classification within the Juglandaceae. Ann Mo Bot Gard 65:1058–1087

Matthies D, Bräuer I, Maibom W, Tscharntke T (2004) Population size and the risk of local extinction: empirical evidence from rare plants. Oikos 105:481–488

McGranahan G, Leslie C (1991) Walnuts (Juglans). Genet Resour Temp Fruit Nut Crops 290:907–974

Millar MA, Byrne M, Nuberg IK, Sedgley M (2012) High levels of genetic contamination in remnant populations of *Acacia saligna* from a genetically divergent planted stand. Restor Ecol 20:260–267

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Syst Biol 51:238–254

Mousset S, Derome N, Veuille M (2004) A test of neutrality and constant population size based on the mismatch distribution. Mol Biol Evol 21:724–731

Mu YL, Xi RT, Lu ZR (1990) Microsporogenesis observation and karyotype analysis of some species in genus *Juglans* L. J Wuhan Bot Res 3:301–310 **(In Chinese)**

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

Pei D, Li R, Liu Z, Wang F (2006) Conservation, exploration and utilization of *Juglans hopeiensis* Hu resources. For Resour Manag 4:66–69 **(In Chinese)**

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a program for detecting recent effective population size reductions from allele data frequencies. J Hered 90:502–503

Pollegioni P, Woeste K, Olimpieri I, Marandola D, Cannata F, Malvolti ME (2011) Long-term human impacts on genetic structure of Italian walnut inferred by SSR markers. Tree Genet Genomes 7:707–723

Pollegioni P, Woeste KE, Chiocchini F, Del Lungo S, Olimpieri I, Tortolano V, Clark J, Hemery GE, Mapelli S, Malvolti ME (2015) Ancient humans influenced the current spatial genetic structure of common walnut populations in Asia. PLoS One 10:e0135980

Pollegioni P, Woeste K, Chiocchini F, Del Lungo S, Ciolfi M, Olimpieri I, Tortolano V, Clark J, Hemery GE, Mapelli S, Malvolti ME (2017) Rethinking the history of common walnut (*Juglans regia* L.) in Europe: its origins and human interactions. PloS ONE 12:e0172541

Poudel RC, Möller M, Li DZ, Shah A, Gao LM (2014) Genetic diversity, demographical history and conservation aspects of the endangered yew tree Taxus contorta (syn. *Taxus fuana*) in Pakistan. Tree Genet Genomes 10:653–665

Pouget M, Youssef S, Migliore J, Juin M, Médail F, Baumel A (2013) Phylogeography sheds light on the central–marginal hypothesis in a Mediterranean narrow endemic plant. Ann Bot 112(7):1409–1420

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

Raymond M, Rousset F (1995) GENEPOP, version 1.2: polulation genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Ruan Y, Huang BH, Lai SJ, Wan YT, Li JQ, Huang S, Liao PC (2013) Population genetic structure, local adaptation, and conservation genetics of *Kandelia obovata*. Tree Genet Genomes 9:913–925

Shu Z, Zhang X, Yu D, Xue S, Wang H (2016) Natural hybridization between Persian walnut and Chinese walnut revealed by simple sequence repeat markers. J Am Soc Sci 141:146–150

Stanford AM, Harden R, Parks CR (2000) Phylogeny and biogeography of *Juglans* (Juglandaceae) based on *matK* and *ITS* sequence data. Am J Bot 87:872–882

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162–1169

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

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting and genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Victory ER, Glaubitz JC, Rhodes OE Jr, Woeste KE (2006) Genetic homogeneity in *Juglans nigra* (Juglandaceae) at nuclear microsatellites. Am J Bot 93:118–126

Vilà M, Weber E, Antonio CM (2000) Conservation implications of invasion by plant hybridization. Biol Invasions 2:207–217

Wang H, Zhang Z, Wang W, Gao Y, Sun H (2005) Studies on photosynthetic characteristics of *Juglans hopeiensis* Hu in the field. Actahorticulturae Sin 32:392–396 **(In Chinese)**

Wang H, Pan G, Ma Q, Zhang J, Pei D (2015) The genetic diversity and introgression of *Juglans regia* and *Juglans sigillata* in Tibet as revealed by SSR markers. Tree Genet Genomes 11:1–11

Wang WT, Xu B, Zhang DY, Bai WN (2016) Phylogeography of postglacial range expansion in *Juglans mandshurica* (Juglandaceae) reveals no evidence of bottleneck, loss of genetic diversity, or isolation by distance in the leading-edge populations. Mol Phylogenet Evol 102:255–264

Wu Y, Pei D, Xi K (1999) Analysis of the origin and the taxonomic position of *Juglans hopeiensis* using RAPD markers. Sci Silvae Sin 35:25–30 **(In Chinese)**

Xi RT (1990) Discussion on the origin of walnut in china. Acta Horticulturae (In Chinese)

Zhang JT, Shao D (2015) Attributes of forest diversity in the Yunmeng Mountain National Forest Park in Beijing, China. Appl Ecol Environ Res 13:769–782

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. Mol Ecol 14:3513–3524

Zhao P, Woeste KE (2011) DNA markers identify hybrids between butternut (*Juglans cinerea* L.), and Japanese walnut (*Juglans ailantifolia* Carr.) Tree Genet Genomes 7:511–533

Zhao P, Zhou HJ, Liu ZL, Woeste KE, Hu DF, Dang M, Li ZH, Wang ML, Zhao GF (2014) A review of research progress on molecular phylogeny and biogeography in *Juglans*. Sci Silvae Sin 50:147–157 **(In Chinese)**