Pak. J. Bot., 49(1): 215-225, 2017.

COMPLEX PHYLOGENETIC PLACEMENT OF *ILEX* SPECIES (AQUIFOLIACEAE): A CASE STUDY OF MOLECULAR PHYLOGENY

FAN YI[1,2], LE SUN[1,2], DA-CHENG HAO[3], YONG PENG[1,2,*] FENG-MING HAN[4,*] AND PEI-GEN XIAO[1,2]

[1]*Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Beijing 100193, China*
[2]*Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Beijing 100193, China*
[3]*Biotechnology Institute, School of Environment and Chemical Engineering, Dalian Jiaotong University, Dalian 116028, China*
[4]*Beijing Biomarker Technologies Co., LTD, Beijing 101300, China*
**Corresponding author's email: ypeng@implad.ac.cn; fantasyee8991@163.com*

Abstract

To investigate the phylogenetic relationships among *Ilex* species distributed in China, we analyzed two alignments including 4,698 characters corresponding to six plastid sequences (*matK, rbcL, atpB-rbcL, trnL-F, psbA-trnH*, and *rpl32-trnL*) and 1,748 characters corresponding to two nuclear sequences (ITS and *nepGS*). Using different partitioning strategies and approaches (i.e., Bayesian inference, maximum likelihood, and maximum parsimony) for phylogeny reconstruction, different topologies and clade supports were determined. A total of 18 *Ilex* species was divided into two major groups (group I and II) in both plastid and nuclear phylogenies with some incongruences. Potential hybridization events may account, in part, for those phylogenetic uncertainties. The analyses, together with previously identified sequences, indicated that all 18 species were recovered within Eurasia or Asia/North America groups based on plastid data. Meanwhile, the species in group II in the nuclear phylogeny were placed in the Aquifolium clade, as inferred from traditional classification, whereas the species in group I belonged to several other clades. The divergence time of most of the 18 *Ilex* species was estimated to be not more than 10 million years ago. Based on the results of this study, we concluded that paleogeographical events and past climate changes during the same period might have played important roles in these diversifications.

Key words: Dating, *Ilex*, Molecular phylogeny, Phylogeography.

Introduction

The genus *Ilex* L., which belongs to the holly family Aquifoliaceae Bartl., comprises deciduous and evergreen bushes or trees distributed throughout temperate and tropical regions worldwide (Galle, 1997), many of which have health-promoting and/or therapeutic values in traditional Chinese medicine and many other ethnomedicine (Hao *et al.*, 2013b). Most of these species in Australia, Europe, and Africa became extinct (Manen, 2004), while *Ilex* continues to exist with high diversity in East Asia, North America, and South America (Manen *et al.*, 2010). This genus is also well-represented in China. Two hundred and four species, of which 149 species are endemic, are identified and are distributing in south regions of the Yangtze River and Qinling Mountains, mainly in South and South-west China (www.eFloras.org).

The oldest pollen record of the *Ilex* species dates back to the Turonian or earliest upper Cretaceous period in Australia, approximately 90 million years ago (mya) (Martin, 1977; Savaroglu *et al.*, 2016; Rasheed *et al.*, 2016), but the lineage is thought to be cosmopolitan by the late Cretaceous period, at about 70–85 mya (Lobreau-Callen, 1975; Manen *et al.*, 2002; Muller, 1981). However, theses fossil records remain controversial (Manen *et al.*, 2002).

Although the first worldwide investigation of morphological taxonomy was already conducted over a century ago (Loesener, 1901), the relationships among various *Ilex* species have long been disputed. Baas (1973; 1975) concluded that *Ilex* species could not be distinguished according to wood and leaf anatomy. Later, Lobreau-Callen (1975) found that *Ilex* species could not be separated based on pollen characteristics. Loizeau *et al*. (2005) recently found the floral morphologies of the *Ilex* species are uniform among species but leaf morphologies show high variability within the species, making it difficult to discriminate between different species. Many phylogenetic studies on the genus *Ilex* have shown that the plastid and nuclear phylogenies are often incongruent based on plastid markers (*psbA-trnH*, *rbcL*, *atpB-rbcL*, and *trnL-F*), nuclear markers [the nuclear ribosomal internal transcribed spacer (ITS), the 5S spacer, and the nuclear encoded plastid glutamine synthetase gene (*nepGS*)], as well as amplified fragment length polymorphism (AFLP) analyses (Cuénoud *et al*., 2000; Gottlieb *et al*., 2005; Selbach-Schnadelbach *et al*., 2009; Setoguchi & Watanabe, 2000; Munir, *et al*., 2015; Gürkanli *et al*., 2016). For example, the plastid phylogenies determined using the plastid RFLPs and *trnL-trnF* spacer were found to agree well with the morphological taxonomy for the *Ilex* species in the Bonin Islands and the Ryukyu Islands of Japan, whereas the ITS phylogenies were unrelated (Setoguchi & Watanabe, 2000). Additionally, within two distantly related sympatric species collected in Tenerife (Canary Island), *I. perado* and *I. canariensis*, the plastid markers were also found to correlate well with species identity (Manen, 2004). However, a recent study showed that the nuclear phylogeny tree was more compatible with current taxonomic classification and the plastid phylogeny was strongly correlated with the geographic distribution (Manen *et al.*, 2010). Therefore, Manen *et al*. (2002) concluded that the *Ilex* species most likely experienced a more complex evolutionary path.

In the present study, we analyzed the phylogenetic relationships among the 18 *Ilex* species collected in a small region of China, mainly from Hunan Province of central south China (Fig. 1), using different partitioning strategies and approaches. Both plastid markers (*matK*, *rbcL*, *trnL-F*, *atpB-rbcL*, *psbA-trnH*, and *rpl32-trnL*) and nuclear markers (ITS and *nepGS*) were used in the analysis. Among them the *rpl32-trnL* sequence was first used in the phylogenetic reconstruction of *Ilex* species (Deng *et al*., 2015). The *rpl32-trnL* spacer is located in the small single-copy (SSC) region of the plastid genome (Shaw *et al*., 2007), with an average length of 1,018 bp, ranging from 543–1,417 bp in different species. This spacer has recently been widely used in phylogenetic analyses (Calvente *et al*., 2011; Gutiérrez-Rodríguez *et al*., 2011; Zecca *et al*., 2012). In combination of additional published molecular data, the phylogenetic positions of the 18 species in the *Ilex* genus were proposed. Finally, we examined the divergence times and causes of diversification for the 18 *Ilex* species.

Materials and Methods

Taxon: Eighteen *Ilex* species (19 samples) and three outgroup taxa, *Helwingia japonica*, *Phyllonoma ruscifolia*, and *P. laticuspis*, were included in the study. Outgroups were selected based on published data (Manen, 2004; Manen *et al*., 2010). Detailed information about the taxon is provided in Table 2. In addition, the molecular data of 108 species (116 samples) described by Manen *et al*. (2010) were used to investigate the phylogenetic positions of the 18 Chinese species within the *Ilex* genus.

DNA extraction, marker amplification, sequencing, and sequence alignment: Genomic DNA was extracted using the Plant Genomic DNA Kit (TianGen, Beijing, China) according to the manufacturer's instructions. Polymerase chain reaction (PCR) was performed using an MJ Research PTC-200 Peltier Thermal Cycler (MJ Research, Watertown, MA), and amplifications were conducted in a 50 μL reaction system containing 15 ng genomic DNA, 5 pmol each primer (for sequences see Table 3), 0.2 mM dNTPs, 2 mM $MgCl_2$, and 1 U DNA polymerase (Takara in Beijing, China). Briefly, after a denaturation step at 94°C for 2 min, amplification was carried out with 36 cycles at a melting temperature of 94°C for 30 sec, an annealing temperature of 50–55°C for 30 sec, and an extension temperature of 72°C for 1 min. Finally, an extra extension step at 72°C for 10 min was included. For sequences that were difficult to amplify using the normal PCR procedure, a "slow and cold" program described by Shaw *et al*. (2005) was used.

Six plastid regions (*matK*, *rbcL*, *atpB-rbcL*, *trnL-F*, *psbA-trnH*, and *rpl32-trnL*) and two nuclear regions (ITS and *nepGS*) were amplified. If two ITS bands appeared, the longer band was used in the analysis. The ITS region includes the partial internally transcribed spacer 1 (ITS1), 5S rRNA gene, and partial internally transcribed spacer 2 (ITS2). For the *trnL-F* region, two primer pairs [primers "c", "d" and primers "e", "f" of Taberlet *et al*. (1991) were used to amplify the *trnL* intron and *trnL-trnF* spacer, individually. All PCR products were sequenced by Sangon Inc. (Shanghai, China). To ensure sequence quality, each sample was bi-directionally sequenced at least three times. Sequence assembly was performed with SeqMan II (DNASTAR, Inc., Madison, WI). The sequences were aligned with ClustalX (Thompson *et al*., 2002), and then they were manually adjusted using BioEdit 7.1.3 (Hall, 1999). All sequences obtained in the present study have been deposited in NCBI Genbank (Table 2).

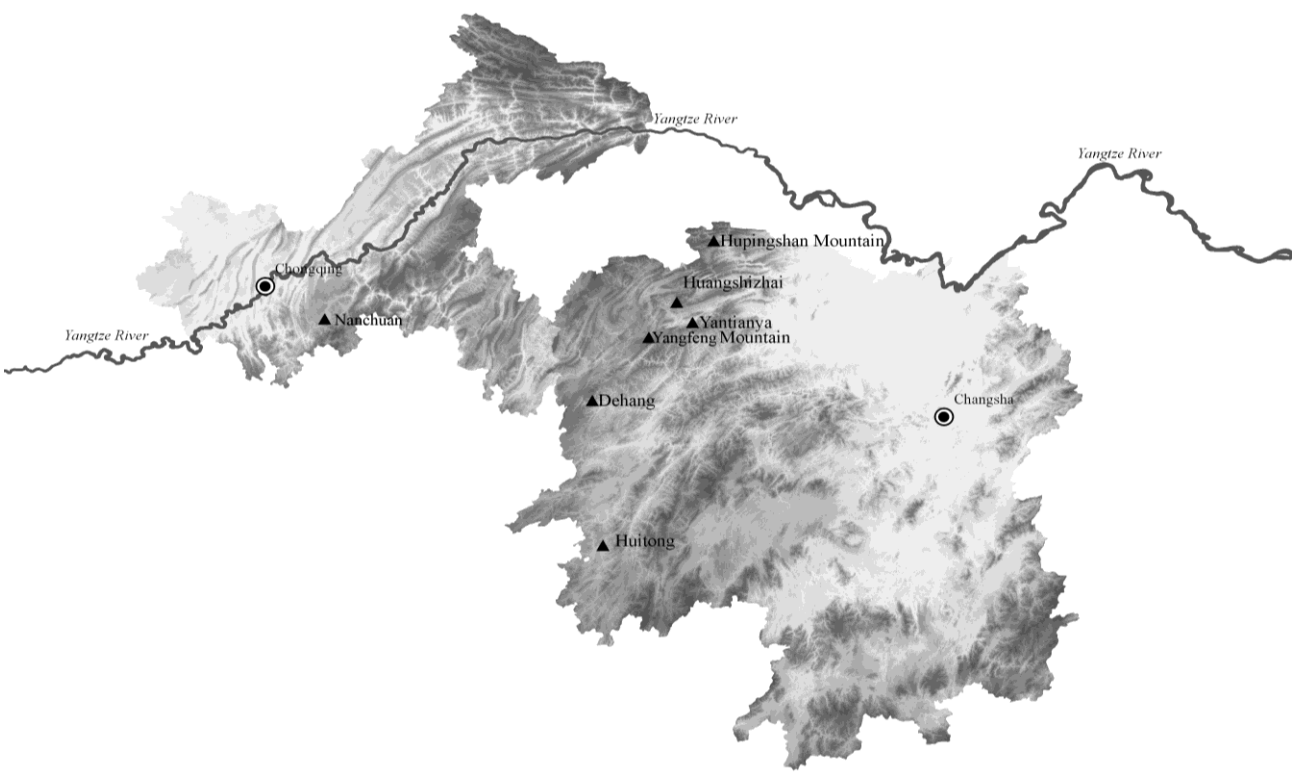

Fig. 1. Collecting areas for this study are indicated as black triangle on the map.

Sequence characterization: The homogeneity of the base compositions among the taxa was examined for possible sources of discordance in phylogenetic signals among loci using the Chi-square test implemented in PAUP*4.0b10 (Swofford, 2003). The program DAMBE (Xia & Xie, 2001) was used to test the nucleotide substitution saturation based on the I_{ss} statistic for the dataset. Congruence between the different partitions was evaluated using the incongruence length difference (ILD) test (Farris *et al*., 1994) as implemented in PAUP* with 1,000 partition replicates, each with 10 random sequence addition replicates saving no more than 500 trees per replicate and Tree Bisection and Reconnection branch swapping.

Phylogenetic analyses: Most parsimonious (MP) trees were sought using the parsimony ratchet (Nixon, 1999). Ten parsimony ratchet searches were conducted, utilizing the PAUPRat program of Sikes & Lewis (2001). Each ratchet search used 1,000 replicates. Then the shortest trees were filtered and used to obtain the majority-rule consensus tree in PAUP*. Support for individual branches was evaluated using non-parametric bootstrapping (Felsenstein, 1985), with 1,000 bootstrap pseudoreplicates and 10 random-taxon-addition sequence replicates per bootstrap pseudoreplicate. Maximum likelihood (ML) analysis was conducted using RAxML 7.0 (Stamatakis *et al*., 2008), the model was set to GTRGAMMA for a concatenated sequence or single partition, and the bootstrap analysis was set to 10,000 replicates. The Bayesian inference (BI) analysis was performed with MrBayes 3.2 (Ronquist & Huelsenbeck, 2003), of which a Markov chain Monte Carlo (MCMC) procedure was employed in order to simultaneously estimate an optimal phylogenetic tree and the posterior probabilities of the interior branches (Li *et al*., 2000). Four parallel Markov chains, heated in increments of 0.01°C, were run for 20 million generations, and the sampling was performed every 1,000 trees. The first 25% of generations were discarded as the burn-in period. In BI and ML analyses, the effect of data partitioning was investigated using four differently partitioned strategies. The first involved no partitions (concatenated sequence), the second partitioned the data by loci, the third partitioned the data by gene and translational pattern, and the fourth partitioned the data by gene, translational pattern, and codon position. Finally, jModeltest version 0.1.1 (Posada, 2008) was used to select the best-fit likelihood model for each data partition strategy. The selected models are listed in Table 1.

Estimation of divergence times: Two calibration constraints are available to calibrate trees of the *Ilex* species, as shown by Manen *et al*. (2010). One calibration point, the most recognized fossil of *Ilex* from Central Europe (Knobloch & Mai, 1986; Manen *et al*., 2002), was dated to approximately 69 mya. The other calibration point was dated to 80 mya and represents the most recent common ancestor of *Phyllonoma*, *Helwingia*, and *Ilex*. This was inferred from estimates given by Davies *et al*. (2004) on an angiosperm tree and by Bremer *et al*. (2004) on an Asterid tree.

To estimate approximate times of lineage divergence, we relied on the Bayesian relaxed-clock method of Drummond *et al*. (2006) under the uncorrelated lognormal rate variation model with default parameters as implemented in BEAST software (Drummond & Rambaut, 2007). Two independent analyses were performed for 100 million generations (10% burn-in), and the sampling was performed every 1,000 generations. Tracer v1.4.1 (Rambaut & Drummond, 2007) was used to plot the log-likelihood scores against generation time to evaluate run convergence. The effective sample sizes for all parameters of interest were greater than 200. Finally, the trees were summarized with TreeAnnotator v1.4.8 (Drummond & Rambaut, 2007) to obtain a maximum clade credibility tree with the estimated divergence times.

Table 1. Specific parameters for each gene and the substitution model for the partitions.

Loci	Variable/informative	Substitution model	-lnL
matK codon pos 1th	25/6	TPM1uf	436.8331
matK codon pos 2th	21/5	HKY	415.1293
matK codon pos 3th	36/7	TPM1uf	480.2062
matK	82/18	TPM1uf	1350.8376
rbcL codon pos 1th	15/8	F81+I	428.9606
rbcL codon pos 2th	8/5	F81+I	388.9124
rbcL codon pos 3th	19/4	TPM1uf	455.128
rbcL	42/17	HKY+G	1325.4236
atpB-rbcL spacer	89/26	TVM+G	1711.9179
trnL-trnF spacer	71/13	TPM3uf	913.3907
psbA-trnH spacer	197/72	TrN+G	1624.6466
rpl32-trnL spacer	31/21	TIM1+I	1508.1774
trnL intron	51/12	TIM1	1067.1837
plastid concatenated	563/179	TPM1uf+I+G	10050.2928
ITS	294/102	TIM2+G	3175.4018
nepGS codon pos 1th	16/none	HKY	154.4744
nepGS codon pos 2th	3/none	F81	104.2061
nepGS codon pos 3th	3/1	TPM1	107.072
nepGS introns	255/17	TPM1uf+I	1937.3188
nepGS exons	22/1	JC	394.6227
nepGS	277/18	HKY	2388.6747
nuclear concatenated	571/120	TIM1+G	5685.7657

Table 2. List of the species, GenBank accession numbers for each marker.

Sample	Voucher	GenBank Accession Number								
		ITS	*nepGS*	*matK*	*rbcL*	*atpB-rbcL*	*trnL-trnF*	*psbA-trnH*	*rpl32-trnL*	*trnL intron*
Ilex ficoidea	F. Yi	KF255665	-	-	KF255693	KF255710	KF255725	KF255741	KF255759	KF255777
Ilex hirsuta	F. Yi	KF255666	-	-	KF255694	KF255711	KF255726	KF255742	KF255760	KF255778
Ilex hainanensis	F. Yi	KF255667	KF255682	-	KF255695	KF255712	KF255727	KF255743	KF255761	KF255779
Ilex hirsuta	F. Yi	KF255668	-	-	KF255696	KF255713	KF255728	KF255744	KF255762	KF255780
Ilex shennongjiaensis	F. Yi	-	-	-	-	-	-	KF255745	-	-
Ilex pentagona	F. Yi	KF255669	KF255683	-	KF255697	KF255714	KF255729	KF255746	KF255763	-
Ilex centrochinensis	F. Yi	KF255670	KF255684	KF255688	KF255698	KF255715	KF255730	KF255747	KF255764	KF255781
Ilex macrocarpa	F. Yi	KF255671	-	-	KF255699	KF255716	KF255731	KF255748	KF255765	KF255782
Ilex pubescens	F. Yi	KF255672	KF255685	KF255689	KF255700	KF255717	KF255732	KF255749	KF255766	KF255783
Ilex pernyi	F. Yi	KF255673	KF255686	KF255690	KF255701	KF255718	KF255733	KF255750	KF255767	KF255784
Ilex lihuaensis	F. Yi	KF255674	-	KF255691	KF255702	KF255719	KF255734	KF255751	KF255768	KF255785
Ilex corallina var. *aberrans*	F. Yi	KF255675	KF255687	KF255692	KF255703	KF255720	KF255735	KF255752	KF255769	KF255786
Ilex cornuta	F. Yi	-	-	-	KF255704	KF255721	-	-	KF255770	-
Ilex chinensis	F. Yi	KF255676	-	-	KF255705	-	KF255736	KF255753	KF255771	-
unidentified Ilex species	F. Yi	KF255677	-	-	-	-	-	KF255754	KF255772	KF255787
Ilex chieniana	F. Yi	KF255678	-	-	KF255706	-	KF255737	KF255755	KF255773	KF255788
Ilex nanchuanensis	F. Yi	KF255679	-	-	KF255707	KF255722	KF255738	KF255756	KF255774	KF255789
Ilex corallina	F. Yi	KF255680	-	-	KF255708	KF255723	KF255739	KF255757	KF255775	KF255790
Ilex latifrons	F. Yi	KF255681	-	-	KF255709	KF255724	KF255740	KF255758	KF255776	KF255791

Table 3. Primers for PCR amplification of the sampled genes in this study.

Region	Name	Primer sequence (5'-3')	Source
matK	3F_KIM	CGTACAGTACTTTTGTGTTTACGA	Ki-Joong Kim, unpublished
	1R_KIM	ACCCAGTCCATCTGGAAATCTTGGTTC	Ki-Joong Kim, unpublished
rbcL	1F	ATGTCACCACAAACAGAAAC	Fay *et al*., 1997
	724R	TCGCATGATCCTGCAGTAGC	Fay *et al*., 1997
atpB-rbcL	2	GAAGTAGTAGGATTGATTCTC	Manen *et al*., 1994
	5	TACAGTTGTCCATGTACCAG	Manen *et al*., 1994
trnL-trnF	e	GGTTCAAGTCCCTCTATCCC	Taberlet *et al*., 1991
	f	ATI'TGAACTGGTGACACGAG	Taberlet *et al*., 1991
psbA-trnH	psbA	GTTATGCATGAACGTAATGCTC	Sang *et al*., 1997
	trnH	CGCGCATGGTGGATTCACAATCC	Tate & Simpson, 2003
	trnHR	GAACGACGGGAATTGAAC	C. J. Cox, unpublished
rpl32-trnL	rpl32-F	CAGTTCCAAAAAAACGTACTTC	Shaw *et al*., 2007
	trnL	CTGCTTCCTAAGAGCAGCGT	Shaw *et al*., 2007
trnL intron	c	CGAAATCGGTAGACGCTACG	Taberlet *et al*., 1991
	d	GGGGATAGAGGGACTTGAAC	Taberlet *et al*., 1991
ITS	ITS5	GGAAGGAGAAGTCGTAACAAGG	White *et al*., 1990
	ITS4	TCCTCCGCTTATTGATATGC	White *et al*., 1990
nepGS	GScp687f	GATGCTCACTACAAGGCTTG	Emshwiller & Doyle, 1999
	GScp994r	AATGTGCTCTTTGTGGCGAAG	Emshwiller & Doyle, 1999

Results

Sequence characteristics: The sequence characteristics of the six plastid regions (*matK*, *rbcL*, *atpB-rbcL*, *trnL-F*, *psbA-trnH*, and *rpl32-trnL*) and two nuclear regions (ITS and *nepGS*) are presented in Table 1, where the *trnL-F* consists of the *trnL* intron and *trnL-trnF* spacer. The concatenated plastid sequences included a total of 4,698 characters, of which 179 were informative characters with *Helwingia japonica*, *Phyllonoma laticuspis*, and *Phyllonoma ruscifolia* as outgroups. The concatenated nuclear sequences consisted of 1,748 characters, of which 120 were informative characters with *Helwingia japonica* as the only outgroup. Because compositional bias is a common source of error in phylogenetic reconstructions, we first tested the homogeneity of base compositions across species for the two datasets. The results of Chi-squared tests indicated that the plastid sequences were homogenous (χ^2 =38.207, *df*=63, *P*=0.994), but the nuclear sequences were not (χ^2=269.173, *df*=51, *P*<0.001). Because most *nepGS* sequences encountered sequencing failures, ultimately, only seven sequences were included in the analysis. This might explain the base heterogeneity. In separate Chi-squared tests of the ITS and *nepGS* sequences, the base homogeneity was not rejected. Furthermore, using BI and ML approaches, the congruent topologies for ITS and *nepGS* were identified. Thus, we concluded that the concatenated nuclear sequences did not contain compositional signals that would bias the phylogeny reconstructions. Furthermore, none of the DAMBE tests yielded an observed index of substitution saturation (I_{ss}) greater than the critical one (I_{ssc}), indicating that there was no substitution saturation issue in the sequences. In the ILD test, no significant phylogenetic incongruence was found among the data partitions (*P*>0.1 for all pairs of partitions), indicating that all data partitions could be analyzed simultaneously in phylogeny reconstruction. The plastid and nuclear sequences were analyzed separately in subsequent analysis because remarkable incongruences between them were observed in previous studies (Manen *et al.*, 2010; Manen *et al.*, 2002; Setoguchi & Watanabe, 2000).

Significant variations in DNA sequences between species and fewer variations within species are prerequisites for an ideal DNA barcoding marker to identify various species. Six metrics were employed to characterize inter-specific variation; an excellent barcode should possess a high inter-specific divergence to distinguish different species. Although we cannot calculate intraspecific distance, we are able to examine the distribution of mean and smallest interspecific distances. The results revealed that the ITS region might possess the highest identification efficiency, followed by *psbA-trnH*. Both these two regions significantly higher levels of inter-specific than *atpB-rbcL*, *rbcL*, *rpl32-trnL* and *trnL intron*. The lowest divergence between conspecific individuals, as determined by all inter-specific calculations was exhibited by *trnL-trnF*. Unfortunately, *matK* and *nepGS* have no calculated data.

Phylogenetic inferences

Plastid sequences: Overall, BI analyses resolved all the branches with much better support [posterior probabilities (PPs>0.90)] under at least one partition strategy, except for clades (*I. shennongjiaensis*, *I.chieniana*) (PPs<0.80), and (*I. centrochinensis*, *I. corallina*) (PPs<0.50), than MP and ML analyses (Fig. 2). Different partitioning strategies for phylogeny reconstructions have been found to produce different levels of bootstrap support (BS) or PPs. For example, with the fourth partition strategy, outgroups had a weak support (PP=0.66) with ingroups in BI analysis instead of the high support (PPs>0.90) under the other partition strategies. Based on the different partition strategies, different topologies for plastid sequences also appeared in the BI, ML, and MP analyses. In MP and all the concatenated ML analyses, *Ilex* species were divided into two groups (BS>90% for MP analysis) (Fig. 2), whereas in separately partitioned ML and all BI analyses, the species *I. ficoidea*, *I. macrocarpa*, and *I. lihuaensis* become the base clade of the other *Ilex* species with good support (BS>80%, and PP=1.00, respectively) The rest of the *Ilex* species formed two groups with weak to moderate support (BS=57%–81%) in partitioned ML analyses and with good support (PPs≥0.99) in all BI analyses. One exception was *I. chinensis*, which clustered with group I in concatenated ML and MP analyses but clustered with group

II in all BI analyses. In addition, the level of diversification within group I was higher than that within group II.

Nuclear sequences: All the nuclear phylogenies were recovered with similar topologies in MP, ML, and most of BI analyses, when the results from BI analyses in the second partitioned strategy were excluded. The species used in the nuclear phylogeny reconstruction could be clearly divided into two major groups (BS>60% and PPs>0.80) (Fig. 3) as in the plastid phylogenies: namely group I, consisting of *I. ficoidea*, *I. macrocarpa*, *I. lihuaensis*, *I. hirsuta*, *I. hirsuta*, and *I. nanchuanensis*; and group II, consisting of *I. centrochinensis*, *I. pernyi*, *I. chinensis*, *I. corallina*, *I. pentagona*, *I. latifrons*, *I. hainanensis*, an unidentified *Ilex* species, and *I. chieniana*. However, incongruences between the plastid and nuclear phylogenies also occurred. For example, the *I. corallina* var. aberrans was nested within different groups in the two data sets. Moreover, under the second partitioned BI analysis, *I. pubescens* was clustered into group I, but group I and group II received only weak support (PP=0.69 and PP=0.63, respectively).

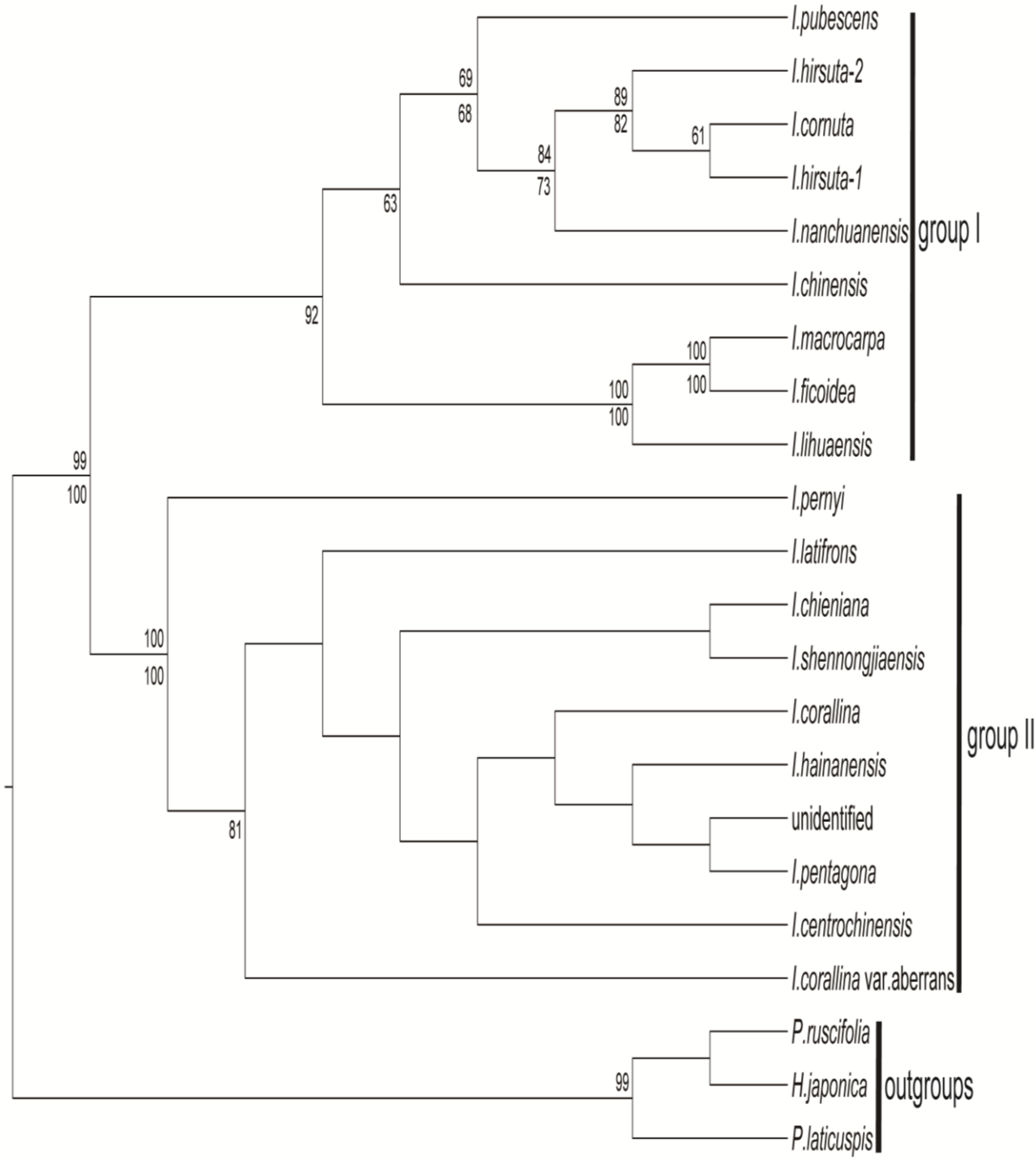

Fig. 2. Plastid phylogeny of *Ilex* recovered under all concatenated maximum likelihood and maximum parsimony methods. The bootstrap values for each node (the highest values for every approach) are shown above and below the branches for Maximum Likelihood and Maximum Parsimony approaches respectively. Only bootstrap values greater than 60% are displayed.

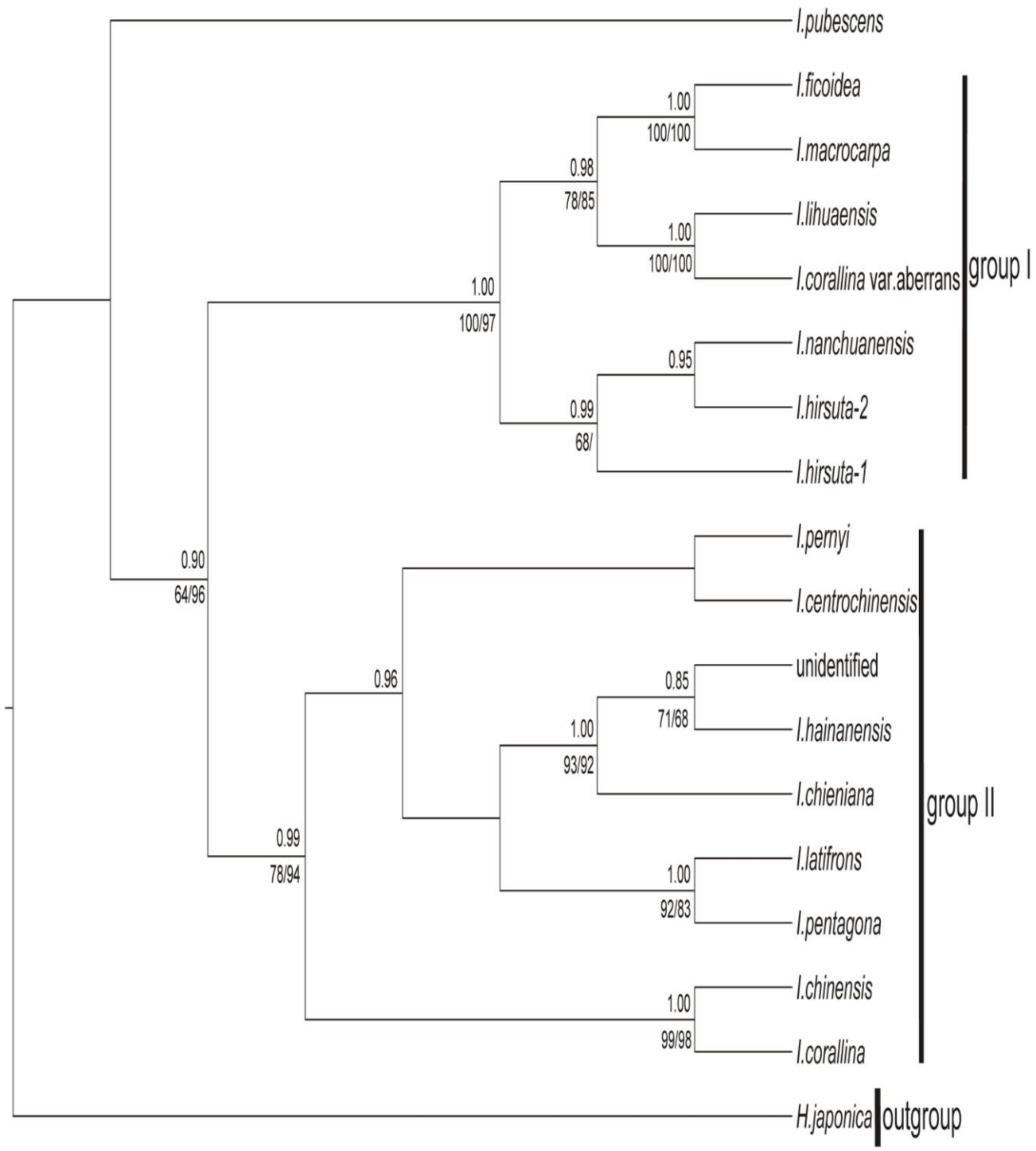

Fig. 3. Nuclear phylogeny of *Ilex* obtained using Bayesian, maximum likelihood, and maximum parsimony approaches (except for the Bayesian approach under the second partitioned strategy). Posterior probabilities and bootstrap values for each node (the highest values for every approach) are shown above and below the branches for Bayesian, maximum likelihood, and maximum parsimony approaches, respectively. Only posterior probabilities greater than 0.80 and bootstrap values greater than 60% are displayed. The columns refer to the two major groups and outgroup.

Phylogenetic positions in Ilex species: When we added the 108 species to our phylogenetic trees, incongruences also occurred between phylogeny, but the phylogenetic positions of the major groups remained stable. Group I (Fig. 2) was clustered with the Asia/North America clades, and group II (Fig. 2) was clustered with the Eurasia clade in plastid trees. In nuclear trees, all of group II (Fig. 3) was incorporated into the Aquifolium clade obtained based on traditional classification; however, group I (Fig. 3) was nested within multiple other clades (Fig. 4). In addition, all these phylogenies were recovered with higher support in terminal clades than in basal clades. It was particularly striking that the same species from different geographic areas could show a relatively distant phylogenetic relationship. For example, the two *I. pubescens* samples were nested within the Stigmatophore clade (our sample) and Rotunda alliance clade (sample given by Manen *et al*., 2010) in nuclear phylogeny (Fig. 4).

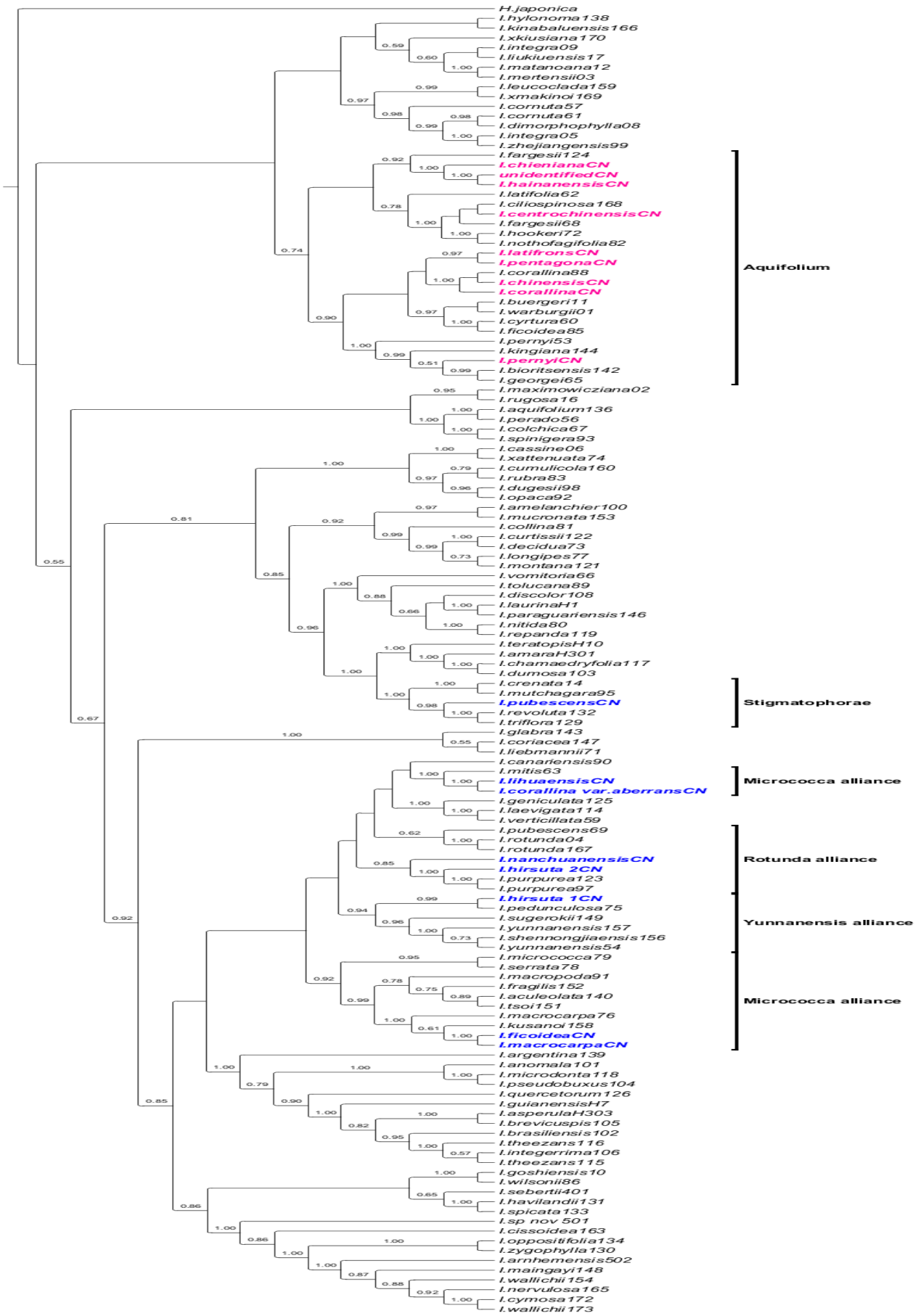

Fig. 4. Nuclear phylogeny of *Ilex* obtained using Bayesian inference. Posterior probabilities for each node are shown above the branches. The group I and group II in Fig. 3 are highlighted in blue and red, respectively. The outgroups are according to Manen *et al*. (2010).

Divergence time: The mean Bayesian divergence time estimates indicated that most of our *Ilex* species diversified from the Middle Miocene period, about 10 mya to present (<1 mya). The divergence times of some species in group I were more ancient based on nuclear data. The most recent common ancestor (MRCA) of group II diversified during the Miocene period, about 15 mya, and this divergence time was supported by plastid and nuclear data.

Discussion

Phylogenetics of Ilex: In the present study, several well-supported contradictory relationships were found between the plastid and nuclear topologies. In addition, the same species in different geographic areas often recovered distinct phylogenetic positions, especially in case of *I. pubescens*. Furthermore, some inconsistencies were found between our results and those of previous studies. For example, *I. shennongjiaensis* was located at the terminal clade comprising *I. shennongjiaensis* and *I. chieniana* in our plastid phylogeny analysis, similar to what has been proposed using both plastid data (Cuénoud *et al*., 2000) and nuclear data (Manen *et al*., 2010). However, *I. shennongjiaensis* has been placed at the basal clade comprising *I. shennongjiaensis* and *I. perado* in a morphology-based phylogenetic tree (Loizeau *et al.*, 2005). Incongruences may have resulted from the limited sampling and nucleotide compositions. Zhang *et al*. (2010) found that sample size can potentially bias genetic diversity and species identification. Likewise, nucleotide bias, character dependence, and character choice can also make the positions of lineages in a phylogenetic tree unstable (Cameron *et al.*, 2004; Dell'Ampio *et al.*, 2009). In our study, some incongruences such as those from positions of lineages or support probabilities indeed occurred among different approaches and different partitioning strategies. Moreover, a complex evolutionary history (such as incomplete lineage sorting, hybridization, and introgression) can also explain this situation. Past research has shown that potential hybridization events occurred in the *Ilex* species (Hao *et al.*, 2013c; Manen, 2004; Manen *et al*., 2010; Setoguchi & Watanabe, 2000; Hussain *et al*., 2016) *I. latifolia*, *I. kudingcha*, *I. cornuta* and *I. pentagona* intermingled in a cluster and these species could not be discriminated in the chloroplast phylogenetic tree. These recurrent hybridization processes could result in complex phylogenetic placement, which in turn increased incongruences among *Ilex* species. In addition, it is believed that plants in similar habitats located at distant regions may exhibit similar morphologies due to adaptation to similar environments, which influences the consistency between molecular and morphological phylogenies (Clark *et al.*, 2012; Loizeau *et al*., 2005; Qiu *et al*., 1995; Schlumpberger & Renner, 2012). Thus, it was not surprising that discordances between phylogenetic and morphological patterns were frequently found in the *Ilex* species because of their worldwide distribution.

Numerous metabolites including saponins and phenylpropanoids have been reported from *Ilex* species. According to the chemotaxonomy, two species have similar chemical constituents which means they may have a closer relationship in phylogenetics. Based on our molecular-level study, *I. latifolia*, *I.cornuta* and *I. pentagona* appear to have a close relationship. This is consistent with the conclusion from Hao *et al.* (2013b) that *I.kudingcha*, *I. cornuta* and *I. pentagona* intermingled in a cluster and were not discriminated these species in the chloroplast phylogenetic tree. Polyacrylamide gel electrophoresis analysis of band differences of esterase isoenzymes from five Aquifoliaceae species suggested that *I. latifolia* should be closer to *I. pentagona* when compared to *I. kudingcha*, while *I. corunta* should be further to *I. latifolia* when compared to *I. kudingcha* (Hao *et al.*, 2013b). This appeared to agrees with our results. For other rare Aquifoliaceae species such as *I.lihuaensis*, *I. shennongjia*, *I.chieniana* and *I.hirsute*, there is little chemotaxonomical information. It would be intriguing to confirm the relationship between DNA barcoding taxology and chemotaxonomy of these species.

Dating and Phylogeography: Although it is unpractical to reconstruct a complete evolutionary scenario within existing *Ilex* species in China without greater numbers of samples and fossil records from throughout China, the results from our present study may provide some clues regarding the history of *Ilex* in China. The divergence ages were quite young, and most of our species were formed no more than 10 mya, ranging from the Miocene epoch, through the Quaternary period to the present day. By tracing the paleogeographical events of China during this period, it was found that the uplift of the Qinghai-Tibetan Plateau (QTP), happened from the late Tertiary/mid-Miocene epoch [~15-10 mya; (Royden *et al*., 2008)], throughout the late Pliocene epoch (~3 mya) and up to Quaternary period. This could dramatically reshape the topographies of western and southern China where most endemic *Ilex* species grow today. If the gene flow between *Ilex* species was blocked by mountains, rivers, and other geographical structure resulting in isolation of their habitats, allopatric species in the *Ilex* genus might have arisen. This, along with rapid species radiation, is also found in many other medicinal genera, e.g., *Clematis*, *Aconitum* and *Fritillaria* (Hao *et al*., 2013a). Also, climatic change is thought to be another main driver of species diversification. This is because plants would be subjected to different selection and new endemic species might occur when the cooling temperature forced the plants to recede some refugia. Once the temperature increased, the allopatric species from different isolated areas could make contact again, leading to hybridization under certain conditions. Previous studies have shown that the climate began to cool down from the Oligocene epoch (~35mya) to very low temperatures through the Pleistocene epoch worldwide (Hewitt, 1999; Webb & Bartlein, 1992). Thus, it is possible that the *Ilex* species have diverged at a high level since the Miocene epoch.

Conclusions

It appears that the analyses combined with previously identified sequences, based on plastid data, indicated that all 18 species were recovered within Eurasia or Asia/North America groups. And based on nuclear data, the group II species were placed in the Aquifolium clade which were matched the traditional classification,

whereas the group I species excluded from this clade. We estimated the divergence time of these 18 *Ilex* was to be not more than 10 million years ago. In general, we concluded that the paleogeographical events and past climate changes may responsible for these diversifications within the genus *Ilex*.

Acknowledgments

This research was supported by the National Natural Science Foundation of P. R. China (No. 81072995, No. 81274188). We are very grateful to Dr. Jin-feng Chen (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences) for his valuable scripts to phylogeny construction. We would like to thank Dr. Jun-jie Zuo (School of Art Education, Jiangsu University) for assistance with mapping drawing.

References

Baas, P. 1973. Thewood anatomical range in *Ilex* (Aquifoliaceae). *Blumea*, 21(2): 193-258.

Baas, P. 1975. Sphenostemon, Phelline, and Oncotheca. *Blumea*, 22(3): 311-407.

Bremer, K. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Syst. Biol.*, 53(3): 496-505.

Calvente, A., D.C. Zappi, F. Forest and L.G. Lohmann. 2011. Molecular phylogeny of *tribe Rhipsalideae* (Cactaceae) and taxonomic implications for *Schlumbergera* and *Hatiora*. *Mol. Phylogenet. Evol.*, 58(3): 456-468.

Cameron, S.L., K.B. Miller, C.A. D'Haese, M.F. Whiting and S.C. Barker. 2004. Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea *sensu lato* (Arthropoda). *Cladistics*, 20(6): 534-557.

Clark, J.L., M.M. Funke, A.M. Duffy and J.F. Smith. 2012. Phylogeny of a Neotropical clade in the Gesneriaceae: more tales of convergent evolution. *Int. J. Plant. Sci.*, 173(8): 894-916.

Cuénoud, P., M. Martinez, P.A. Loizeau, R. Spichiger, S. Andrews and J.F. Manen. 2000. Molecular phylogeny and biogeography of the genus *Ilex* L. (Aquifoliaceae). *Ann. Bot.*, 85(1): 111-122.

Davies, T.J., T.G. Barraclough, M.W. Chase, P.S. Soltis, D.E. Soltis and V. Savolainen. 2004. Darwin's abominable mystery: insights from a supertree of the angiosperms. *P. Natl. Acad. Sci. USA.*, 101(7): 1904-1909.

Dell'Ampio, E., N.U. Szucsich, A. Carapelli, F. Frati, G. Steiner, A. Steinacher and G. Pass. 2009. Testing for misleading effects in the phylogenetic reconstruction of ancient lineages of hexapods: influence of character dependence and character choice in analyses of 28S rRNA sequences. *Zool. Scr.*, 38(2): 155-170.

Deng, J.B., J. Liu, K.S. Ahmad, C.B. Ding, L. Zhang, Y.H. Zhou and R.W. Yang. 2015. Relationships evaluation on six herbal species (*Curcuma*) by DNA barcoding. *Pak. J. Bot.*, 47(3): 1103-1109.

Drummond, A.J. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.*, 7(214): 1-8.

Drummond, A.J., S.Y. Ho, M.J. Phillips and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PloS. Biol.*, 4(5): 699-710.

Farris, J.S., M. Källersjö, A.G. Kluge and C. Bult. 1994. Testing significance of incongruence. *Cladistics*, 10(3): 315-319.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4): 783-791.

Galle, F.C. 1997. *Hollies: the genus Ilex*. Timber Press, Portland.

Gottlieb, A.M., G.C. Giberti and L. Poggio. 2005. Molecular analyses of the genus *Ilex* (Aquifoliaceae) in southern South America, evidence from AFLP and ITS sequence data. *Am. J. Bot.*, 92(2): 352-369.

Gürkanli, C.T., E.B. Aydin, E. Demirci, I. Erper, G. Karaca, T.F. Hsieh, L.S. Poltronieri and I. Özkoç. 2016. 18S rDNA and β-tubulin diversity in *Rhizoctonia Zeae* Voorhees. *Pak. J. Bot.*, 48(2): 645-651.

Gutiérrez-Rodríguez, C., J.F. Ornelas and F. Rodríguez-Gómez. 2011. Chloroplast DNA phylogeography of a distylous shrub (*Palicourea padifolia*, Rubiaceae) reveals past fragmentation and demographic expansion in Mexican cloud forests. *Mol. Phylogenet. Evol.*, 61(3): 603-615.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Aci. Symp. Ser.*, 41: 95-98.

Hao, D.C., X.J. Gu, P.G. Xiao and Y. Peng. 2013a. Phytochemical and biological research of Fritillaria medicine resources. *Chin. J. Nat. Med.*, 11(4): 330-344.

Hao, D.C., X.J. Gu, P.G. Xiao, L.J. Xu and Y. Peng. 2013c. Recent advances in the chemical and biological studies of Aconitum pharmaceutical resources. *J. Chin. Pharm. Sci.*, 22(3): 209-221.

Hao, D.C., X.J. Gu, P.G. Xiao, Z. Liang, L.J. Xu and Y. Peng. 2013b. Research progress in the phytochemistry and biology of *Ilex* pharmaceutical resources. *Acta Pharm. Sin. B*, 3(1): 8-19.

Hewitt, G.M. 1999. Post-glacial re-colonization of European biota. *Bio. J. Linn. Soc.*, 68(1): 87-112.

Hussain, I., A. Sulatan, Z.K. Shinwari, G. Raza and K. Ahmed. 2016. Genetic diversity based on morphological traits in walnut (*Juglans regia* L.) Landraces from Karakoram region-1. *Pak. J. Bot.*, 48(2): 653-659.

Knobloch, E. and D.H. Mai. 1986. Monographie der Fruchte und Samen in der Kreide von Mitteleuropa. *Rozpray Ustredniho ustavu geologickeho*, 109: 566.

Li, S., D.K. Pearl and H. Doss. 2000. Phylogenetic tree construction using Markov chain Monte Carlo. *J. Am. Stat. Assoc.*, 95(450): 493-508.

Lobreau-Callen, D. 1975. Les variations dimensionnelles du pollen du genre *Ilex* (Aquifoliaceae) et leurs rapports avec le climat. *B. Soc. Bot. Fr.*, 122(sup1): 179-199.

Loesener, T. 1901. *Monographia aquifoliacearum*. Ehrhardt Karras Press.

Loizeau, P.A., G. Barriera, J.F. Manen and O. Broennimann. 2005. Towards an understanding of the distribution of *Ilex* L.(Aquifoliaceae) on a World-wide scale. *Biol. Skr.*, 55: 501-520.

Manen, J.F. 2004. Are both sympatric species *Ilex perado* and *Ilex canariensis* secretly hybridizing? Indication from nuclear markers collected in Tenerife. *BMC Evol. Biol.*, 4(1): 46.

Manen, J.F., G. Barriera, P.A. Loizeau and Y. Naciri. 2010. The history of extant *Ilex* species (Aquifoliaceae): evidence of hybridization within a Miocene radiation. *Mol. Phylogenet. Evol.*, 57(3): 961-977.

Manen, J.F., M. Boulter and Y. Naciri-Graven. 2002. The complex history of the genus *Ilex* L.(Aquifoliaceae): evidence from the comparison of plastid and nuclear DNA sequences and from fossil data. *Plant Syst. Evol.*, 235(1-4): 79-98.

Martin, H. 1977. The history of *Ilex* (Aquifoliaceae) with special reference to Australia: evidence from pollen. *Aust. J. Bot.*, 25(6): 655-673.

Muller, J. 1981. Fossil pollen records of extant angiosperms. *Bot. Rev.*, 47(1): 1-142.

Munir, U., A. Perveen and S. Qamarunnisa. 2015. The utility of rbcL and matK regions for DNA barcoding analysis of the genus Suaeda (Amaranthaceae) species. *Pak. J. Bot.*, 47(6): 2329-2334.

Nixon, K.C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics*, 15(4): 407-414.

Posada, D. 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.*, 25(7): 1253-1256.

Qiu, Y.L., M.W. Chase and C.R. Parks. 1995. A chloroplast DNA phylogenetic study of the eastern Asia-eastern North America disjunct section *Rytidospermum* of *Magnolia* (Magnoliaceae). *Am. J. Bot.*, 82(12): 1582-1588.

Rambaut, A. and A.J. Drummond. 2007. Tracer version 1.4. Computer program and documentation distributed by the author. *Beast bio.ed.ac.uk/Tracer.*

Rasheed, A.A., A. Perveen, R. Abid and M. Qaiser. 2016. Pollen morphology of the subfamily Arecoideae Griff. (family-Arecaceae) from Pakistan and Kashmir. *Pak. J. Bot.*, 48(3): 1051-1060.

Ronquist, F. and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12): 1572-1574.

Royden, L.H., B.C. Burchfiel and R.D. Van Der Hilst. 2008. The geological evolution of the Tibetan Plateau. *Science*, 321(5892): 1054-1058.

Savaroglu, F., I.P. Erkara and F. Ozcelik. 2016. Spore morphology of some Grimmiaceae Arn. species belonging to moss flora of Turkey. *Pak. J. Bot.*, 48(3): 1045-1050.

Schlumpberger, B.O. and S.S. Renner. 2012. Molecular phylogenetics of *Echinopsis* (Cactaceae): Polyphyly at all levels and convergent evolution of pollination modes and growth forms. *Am. J. Bot.*, 99(8): 1335-1349.

Selbach-Schnadelbach, A., S.S. Cavalli, J.F. Manen, G.C. Coelho and D.T.T. Souza-Chies. 2009. New information for *Ilex* phylogenetics based on the plastid *psbA-trnH* intergenic spacer (Aquifoliaceae). *Bot. J. Linn. Soc.*, 159(1): 182-193.

Setoguchi, H. and I. Watanabe. 2000. Intersectional gene flow between insular endemics of *Ilex* (Aquifoliaceae) on the Bonin Islands and the Ryukyu Islands. *Am. J. Bot.*, 87(6): 793-810.

Shaw, J., E.B. Lickey, E.E. Schilling and R.L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.*, 94(3): 275-288.

Shaw, J., E.B. Lickey, J.T. Beck, S.B. Farmer, W. Liu, J. Miller, K.C. Siripun, C.T. Winder, E.E. Schilling and R.L. Small. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.*, 92(1): 142-166.

Sikes, D. and P.O. Lewis. 2001. PAUPRat: PAUP* implementation of the parsimony ratchet. University of Connecticut, USA

Stamatakis, A., P. Hoover and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.*, 57(5): 758-771.

Swofford, D. 2003. PAUP*: phylogenetic analysis using parsimony, version 4.0 b10.

Taberlet, P., L. Gielly, G. Pautou and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant. Mol. Biol.*, 17(5): 1105-1109.

Thompson, J.D., T. Gibson and D.G. Higgins. 2002. Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinformatics*, 2(3): 2-3.

Webb, T. and P. Bartlein. 1992. Global changes during the last 3 million years: climatic controls and biotic responses. *Annu. Rev. Ecol. Syst.*, 23: 141-173.

Xia, X. and Z. Xie. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *J. Hered.*, 92(4): 371-373.

Zecca, G., J.R. Abbott, W.B. Sun, A. Spada, F. Sala and F. Grassi. 2012. The timing and the mode of evolution of wild grapes (Vitis). *Mol. Phylogenet. Evol.*, 62(2): 736-747.

Zhang, A., L. He, R. Crozier, C. Muster and C.D. Zhu. 2010. Estimating sample sizes for DNA barcoding. *Mol. Phylogenet. Evol.*, 54(3): 1035-1039.

(Received for publication 15 January 2016)