Botanical Journal of the Linnean Society, 2009, **159**, 182–193. With 3 figures

New information for *Ilex* phylogenetics based on the plastid *psbA-trnH* intergenic spacer (Aquifoliaceae)

ALESSANDRA SELBACH-SCHNADELBACH[1]*, SUZANA SMITH CAVALLI[1], JEAN-FRANÇOIS MANEN[2], GERALDO CENI COELHO[3] and TATIANA TEIXEIRA DE SOUZA-CHIES[1,4]

[1]*Programa de Pós-Graduação em Genética e Biologia Molecular – UFRGS, 91501-970, Porto Alegre, RS, Brazil*
[2]*Conservatoire et Jardin Botaniques, Impératrice 1, CH-1292 Chambésy/Genève, Switzerland*
[3]*Departamento de Biologia e Química, UNIJUÍ, Cx. P. 560, 98.700-000, Ijuí, RS, Brazil*
[4]*Departamento de Botânica – UFRGS, 91501-970, Porto Alegre, RS, Brazil*

Received 13 February 2007; accepted for publication 5 June 2008

The aim of the present work was to clarify the origin and phylogenetic position of the species belonging to the genus *Ilex* (Aquifoliaceae), especially the South American species. Phylogenetic relationships of the genus *Ilex* were investigated using the plastid *psbA-trnH* intergenic spacer and parsimony and Bayesian analyses. The *psbA-trnH* intergenic spacer was shown to evolve slowly within *Ilex*, but a major gap present in this region was useful in the phylogenetic study of the genus. To obtain more potentially parsimonious characters, *atpB-rbcL* intergenic spacer data were combined with those for *psbA-trnH*. Many gaps present in the *psbA-trnH* region were useful in the phylogenetic study of the genus *Ilex*. The topology of the trees showed that, in general, the clades are strongly related to geographical areas, a fact especially evident in certain different Asian lineages. © 2009 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2009, **159**, 182–193.

ADDITIONAL KEYWORDS: Bayesian inference – maximum parsimony – molecular phylogeny – *psbA-trnH* intergenic spacer.

INTRODUCTION

The genus *Ilex* L. (Aquifoliaceae) has more than 400 species. *Ilex* species are dioecious trees or shrubs, mostly perennials, with alternate, simple leaves, small and usually unisexual flowers, superior ovary with one ovule in each locule, the absence of nectaries and a fruit drupe (Judd *et al.*, 1999). The inflorescences are thyrses with lateral cymes (Coelho & Mariath, 1996). The chromosome number in most *Ilex* species is $2n = 40$, but some have higher chromosome numbers: $2n = 72$ (*I. verticillata* A.Gray); $2n = 80$ (*I. argentina* Lillo; *I. anomala* Hook. & Arn.); $2n = 120$ (*I. pedunculosa* Miq.) (Barral, Poggio & Giberti, 1995). *Ilex* includes several hollies cultivated for ornamental use. The genus also includes the species *Ilex paraguariensis* A.St.-Hil., popularly known as 'maté' or 'yerba mate', leaves of which are used in the production of a tea-like beverage commonly consumed in South America. Fossil records of about 250 species from all over the world, excluding Antarctica, have been found, and indicate that the ancestors of the genus *Ilex* may have originated in Gondwana (Galle, 1997).

There are three major geographical centres of current diversity for *Ile*x. The richest and primary area is western Asia, including Taiwan, Japan, China, Korea, the Indochina Peninsula and other regions in the area, the second South America and the third the Malay Peninsula and neighbouring archipelagos (Galle, 1997).

Loesener (1942) proposed the first *Ilex* classification based on morphology and geographical distribution. More recent classifications based on inflorescence morphology (Loizeau & Spichiger, 1992; Loizeau, 1994; Coelho & Mariath, 1996) are, however, not in complete agreement with Loesener's system.

*Corresponding author. Current address: Universidade Estadual do Rio Grande do Sul, Rua Oscar Matzembacher, 475, 96760-000, Tapes, RS, Brazil. E-mail: alessandra.schnadelbach@gmail.com

A previous phylogenetic study of the genus *Ilex* based on the sequence of the plastid *atpB-rbcL* intergenic spacer (Cuénoud *et al.*, 2000) proposed that the genus should be organized as four groups, each with distinct geographical or ecological arrangements. The first is exclusively American, the second is Eurasian, the third comprises all deciduous species, and the fourth is Asian/North American. Broadly speaking, this study contradicted the systematic treatment proposed by Loesener. The results indicated that the *Ilex* lineage was already cosmopolitan before the end of the Cretaceous.

A recent phylogenetic study of the genus *Ilex* based on the plastid *atpB-rbcL* and *trnL-trnF* intergenic spacers and the *rbcL* gene and nuclear ribosomal ITS and 5S spacer showed that the plastid-based phylogenetic trees were strongly related to the geographical distribution of extant species, and the nuclear-based phylogenetic trees suggested frequent interlineage hybridizations and lineage sorting. There were two different *Ilex* lineages in the plastid-based American clade showing two different biogeographical relationship patterns in South America: one including North American species and the other including Asian species. Moreover, the genus probably experienced frequent lineage sorting and interlineage hybridization with subsequent nuclear or cytoplasmic introgression (Manen, Boulter & Naciri-Graven, 2002).

Intersectional hybridization and nuclear gene flow between insular endemic species of the genus *Ilex* from Bonin and Ryukyu Islands, Japan, without evidence of plastid DNA gene flow, were detected by Setoguchi & Watanabe (2000) using plastid restriction fragment length polymorphisms, *trnL-trnF* and ITS sequences. The plastid phylogenetic trees were consistent with the morphologically based taxonomy, whereas the nuclear rDNA phylogenetic trees grouped several putatively unrelated endemic species from both islands. In addition, unilateral introgression between the sympatric species *I. perado* Aiton and *I. canariensis* Poir. from Tenerife (Canary Islands) was reported by Manen (2004), with *I. perado* as the male donor.

Of the genome regions that potentially evolve rapidly, the plastid *psbA-trnH* intergenic spacer may be appropriate for molecular phylogenetic studies at a lower taxonomic level (Sang, Crawford & Stuessy, 1997; Hamilton, Braverman & Soria-Hernanz, 2003). Some studies have shown that it can be highly variable when compared with the *trnL-trnF* region used at the infrageneric level (Sang *et al.*, 1997; Klak *et al.*, 2003; Smissen, Breitwieser & Ward, 2004). This region is now widely used in phylogenetic approaches concerning phylogeographical questions and taxonomic circumscriptions based on interspecific relationships (Holderegger & Abbott, 2003; Kim, Lu & Lepschi, 2004; Butterworth & Wallace, 2005; Bruyns, Mapaya & Hedderson, 2006).

DNA sequences from the *psbA-trnH* intergenic spacer were analysed for species belonging to different groups of *Ilex* to elucidate the phylogenetic position of some species. The *atpB-rbcL* intergenic spacer was combined with *psbA-trnH*, aiming to contribute to a more accurate phylogenetic study of this genus.

MATERIAL AND METHODS

Plant material

Leaf samples were collected in the field and dried in silica gel. Some samples employed in this study were the same as those used by Cuénoud *et al.* (2000). The species included in the present work are listed in Table 1.

DNA extraction, DNA amplification and sequencing

DNA was extracted from dried leaves using 2 × CTAB (cetyltrimethylammonium bromide) protocols, adapted to 2 mL microtubes from Doyle & Doyle (1987) or Saghai-Maroof *et al.* (1984).

Amplifications of the *psbA-trnH* intergenic spacer were performed in 50 μL reaction volumes containing 1 × polymerase chain reaction (PCR) buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 1.0 unit *Taq* DNA Polymerase (Amersham Biosciences) and 40 ng of DNA template. The amplification of the *psb*A-*trn*H intergenic spacer was performed with an initial denaturation at 94 °C for 3 min, followed by 35 cycles of 45 s at 94 °C, 1 min at 55 °C (or 50 °C for some species) and 2 min at 72 °C, with a final extension at 72 °C for 5 min. The primers used were those of Sang *et al.* (1997): forward primer, 5′-GTTATGCATGAACGTAATGCTC-3′; reverse primer, 3′-CGCGCATGGTGGATTCACAAATC-5′.

The forward primer was also used for sequencing reactions. For some samples, the reverse primer was used to complete the entire region or to resolve ambiguous sequences. PCR fragments were purified by enzymatic treatment with exonuclease I and shrimp alkaline phosphatase. The sequencing was performed in a Mega Bace 1000 (Amersham Pharmacia Biotech), ABI Prism 310 or ABI Prism 3100 automated DNA sequencer (Perkin-Elmer).

Phylogenetic analyses

In general, only one specimen for each species was sequenced, as it has been shown that different accessions from one species usually have identical *psbA-trnH* intergenic spacer sequences (Sang *et al.*, 1997). DNA sequences were aligned using the multiple progressive alignment procedure of Clustal W (Thompson,

Table 1. Analysed species, collectors, vouchers and GenBank accession numbers for sequence data

	Collector	Voucher	Accession no.
I. aculeolata Nakai	Zhong Shu Hua		EU359388
I. 'Adonis'*	C.-Y. HU	–	EU359323
I. affinis Gardn.	A. Cervi	–	EU359340
I. amara (Vell.) Loes.	R. Harley	26229	EU359360
I. aquifolium L.	–	MNHN	EU359337
I. argentina Lillo	–	INTA 221	EU359339
I. asperula Mart. ex Reissek	G.C. Coelho	HUI4022	EU359310
I. × *attenuata* Ashe	S. Andrews	SA1515	EU359356
I. beanii Rehder	–	Kew 1973-20564	EU359330
I. bioritsensis Hayata	–	Kew 1985-4641	EU359311
I. brasiliensis Loes.	A. Cervi	–	EU359357
I. buergeri Miq.	J.M. Ruter	Ruter N3-7	EU359334
I. canariensis Poir.	S. Andrews	SA1488	EU359361
I. ciliospinosa Loes.	–	Kew 1929-47703	EU359317
I. cissoidea Loes.	S. Andrews	SA860	EU359376
I. colchica Pojark.	–	Kew 1996-1801	EU359314
I. collina Alexander	S. Andrews	SA1479	EU359369
I. conocarpa Reissek	G.C. Coelho	HUI 4015	EU359329
I. corallina Franch.	S. Andrews	SA1517	EU359379
I. coriacea Chapm.	S. Andrews	SA1536	EU359389
I. cornuta Lindl. & Paxton	–	Kew 1986-8420	EU359378
I. crenata Thunb.	–	Kew 1988-626	EU359312
I. cumulicola Small	S. Andrews	SA1537	EU359390
I. decidua Walter	–	Kew 1969-13532	EU359333
I. dimorphophylla Koidz.	–	Kew 1983-2412	EU359315
I. discolor Hemsl.	A.M. Olivo	–	EU359391
I. dugesii Fernald	S. Andrews	SA1497	EU359358
I. dumosa Reissek	G.C.Coelho	HUI 4004	EU359316
I. 'Elegance'†	–	Kew 1970-3044	EU359313
I. fargesii Franch.	S. Andrews	SA1453	EU359381
I. formosana Maxim.	J.M. Ruter	Ruter N8-20	EU359324
I. fragilis Hook.f.	C.J. Wingfield	–	EU359392
I. geniculata Maxim.	S. Andrews	SA1527	EU359397
I. glabra A.Gray	–	Kew 1985-8411	EU359338
I. hookeri King	S. Andrews	SA1464	EU359395
I. hylonoma Hu & Tang	Zhong Shu Hua	–	EU359359
I. integra Thunb.	–	Kew 1986-4763	EU359336
I. integrifolia Hort.	A.Cervi	–	EU359385
I. intricata Hook.f.	S. Andrews	SA1568	EU359382
I. kinabaluensis S.Andrews	S. Andrews	SA1580	EU359371
I. kingiana Cockerell	S. Andrews	SA1494	EU359372
I. kiusiana Hatus.	S. Andrews	SA1552	EU359374
I. kusanoi Hayata	S. Andrews	SA1512	EU359350
I. latifolia Thunb.	–	Kew 1974-2732	EU359318
I. leucoclada Makino	S. Andrews	SA1542	EU359362
I. liebmannii Standl.	S. Andrews	SA1463	EU359363
I. liukiuensis Loes.	H. Nagamasu	5034	EU359380
I. longipes Chapman ex Trelease	–	Kew 1983-3911	EU359320
I. macrocarpa Oliver	J.M. Ruter	Ruter S8-6	EU359345
I. macropoda Miq.	C.-Y. Hu	–	EU359344
I. makinoi Hara	S. Andrews	SA1520	EU359383
I. matanoana Makino	A. Soejima	930210	EU359364
I. montana Torr. & A.Gray	S. Andrews	SA1546	EU359365

Table 1. *Continued*

	Collector	Voucher	Accession no.
I. myrtifolia Walter	J.M. Ruter	Ruter S12-11	EU359325
I. nipponica Makino	H. Nagamasu	5534	EU359367
I. nitida (Vahl) Maxim.	G.J. Breckon	–	EU359393
I. nothofagifolia Kingdon-Ward	S. Andrews	SA1499	EU359368
I. opaca Aiton	–	Kew 1992-1346	EU359322
I. paraguariensis A.St.-Hil.	–	HAS100443	EU359321
I. paraguariensis A.St.-Hil. var. *vestita* Loes.	G.C. Coelho	–	EU359353
I. pedunculosa Miq.	C.-Y. Hu	–	EU359341
I. perado Webb & Berthel.	C.-Y. Hu	–	EU359347
I. pernyi Franch.	–	Kew 1973-20566	EU359319
I. poneantha Koidz.	J.M. Ruter	Ruter S8-7	EU359335
I. pubescens Hook. & Arn.	S. Andrews	SA1459	EU359332
I. purpurea Hassk.	S. Andrews	SA1529	EU359331
I. quercetorum I.M.Johnst.	S. Andrews	SA1531	EU359370
I. repanda Griseb.	S. Andrews	SA1521	EU359373
I. revoluta Stapf	S. Andrews	SA1584	EU359384
I. rivularis Gardn.	G.C. Coelho	HUI 4001	EU359342
I. rubra S. Watson	J.M. Ruter	Ruter S14-8	EU359346
I. rugosa F.Schmidt	H. Nagamasu	–	EU359386
I. serrata Thunb.	C.-Y.Hu	–	EU359343
I. shennongjiaensis T.R.Dudley & S.C.Sun	S. Andrews	SA1530	EU359375
I. spicata Blume	S. Andrews	SA1592	EU359366
I. spinigera Loes.	J.M. Ruter	Ruter N3-5	EU359349
I. subcordata Reissek	A. Cervi	–	EU359328
I. theezans Mart.	–	HAS 100066	EU359352
I. tolucana Hemsl.	S. Andrews	SA1491	EU359394
I. trichothyrsa Loes.	G.C. Coelho	HUI 4005	EU359327
I. triflora Blume	J.M. Ruter	Ruter S10-9	EU359348
I. tsoii Merr. & Chun	S. Andrews	SA1528	EU359377
I. verticillata A.Gray	Griffiths	ICN127560	EU359326
I. warburgii Loes.	H. Nagamasu	5033	EU359396
I. wilsonii Loes.	S. Andrews	SA1500	EU359354
I. zhejiangensis C.J.Tseng	S. Andrews	SA1493	EU359355
I. zygophylla Merr.	S. Andrews	SA1581	EU359351
Helwingia japonica (Thunb.) F.Dietr.‡	–	Kew 1953-24705	EU359387

*I. 'Adonis' = *I.* × 'Nellie Stevens' (putative hybrid) × *I. latifolia*.
†I. 'Elegance' = *I. integra* × *I. pernyi*.
‡Outgroup.

Higgins & Gibson, 1994), and misalignments were corrected manually.

The analyses were carried out with *psbA-trnH* data for 88 *Ilex* specimens, using *Helwingia japonica* as outgroup, and with the combined data from *psbA-trnH* plus *atpB-rbcL* for 73 *Ilex* specimens, with *H. japonica* as outgroup. The *atpB-rbcL* data were obtained from http://www.cjb.unige.ch (Cuénoud *et al.*, 2000).

Maximum parsimony (MP) analyses were performed using PAUP 4.10b (Swofford, 2002) with the heuristic search option and 'tree bisection–reconnection' (TBR) branch-swapping algorithm, random addition sequences. The accelerated transformation (ACCTRAN) optimization was used to infer the branch lengths. The analyses were carried out with equal weight in all positions, and the gaps, coded as binary characters, were added as an additional matrix of 34 characters (34 indels). Support for the clades was estimated by 1000 bootstrap replications with fast heuristic search, performed with PAUP.

MrModeltest v2 (Nylander, 2004) was used to select the evolutionary model to be employed in the Bayesian inference (BI). BI was performed with the software MrBayes: Bayesian Inference of phylogeny, version 2.01 (Huelsenbeck & Ronquist, 2001). Uniform, prior probabilities and a random starting tree were used. The Markov chain Monte-Carlo (MCMC) procedure was run simultaneously and sampled every 100 generations for a total of 1 000 000 generations. The gaps were considered in a binary matrix (Ronquist, Huelsenbeck & Mark, 2005). The majority rule consensus tree was calculated with PAUP 4.10b (Swofford, 2002).

RESULTS

A fragment approximately 400 bp long, corresponding to the *psbA-trnH* intergenic spacer, was amplified by PCR; *I. discolor* Hemsl. and *I. tolucana* Hemsl. had the smallest *psbA-trnH* intergenic spacers (388 bp), and *I. serrata* Thunb. had the longest (479 bp). The *psbA-trnH* intergenic spacer of *H. japonica* (Thunb.) F.Dietr. was 359 bp long.

Numerous indels (insertions/deletions) were observed, but the high degree of nucleotide similarity resulted in a matrix with little alignment ambiguity. The most important indel corresponded to a fragment that varied in length from 31 bp in *H. japonica* to 74 bp in *I. serrata*. This indel was present at different sizes in *I. aculeolata* Nakai, *I. canariensis* Poir., *I. cissoidea* Loes., *I. collina* Alexander, *I. coriacea* Chapm., *I. crenata* Thunb., *I. decidua* Walter, *I. fragilis* Hook.f., *I. geniculata* Maxim., *I. glabra* A.Gray, *I. kusanoi* Hayata, *I. longipes* Chapman ex. Trelease, *I. macropoda* Miq., *I. macrocarpa* Oliver, *I. montana* Torr. & A.Gray, *I. nipponica* Makino, *I. pedunculosa* Miq., *I. poneantha* Koidz., *I. pubescens* Hook. & Arn., *I. purpurea* Hassk., *I. revoluta* Stapf, *I. shennongjiaensis* T.R.Dudley & S.C.Sun, *I. spicata* Blume, *I. tsoii* Merr. & Chun, *I. triflora* Blume, *I. verticillata* A.Gray, *I. zygophylla* Merr and *I. wilsonii* Loes. It was present only in North American and Asian species, *I. canariensis* (from the Canary Islands) and *H. japonica*. Many autapomorphic indels were also observed in various species.

A total of 573 characters was analysed (539 nucleotides and 34 indels) in the aligned matrix, 145 of which were variable and 101 were potentially parsimony informative. Analyses from the *psbA-trnH* intergenic spacer with equal weights resulted in a huge number of most parsimonious trees, and thus the search was limited to 10 000 replications. The research generated 24 710 equally most parsimonious trees, with a length of 428 steps, consistency index (CI) of 0.67, CI excluding uninformative characters of 0.48 and retention index (RI) of 0.75. The consensus tree resulting from the MP analyses was not shown because it agrees with the majority rule consensus tree calculated from BI. The majority rule consensus tree derived from 14 071 trees from the BI analysis based on the *psbA-trnH* intergenic spacer is shown in Figure 1.

The model selected to perform BI based on the *psbA-trnH* intergenic spacer was F81 plus the gamma shape parameter (gamma = 0.3571), $-\ln L$ = 2680.9670. Stationary conditions were reached around generation 71 300; thus, the first 713 trees were eliminated.

Figure 1 (BI) shows that almost all South American species are grouped (South American clade) with *I. brasiliensis* Loes., *I. liebmannii* Standl and *I. quercetorum* I.M.Johnst. (the last two species are from Central America) in the consensus tree (*I. asperula* Mart. ex Reissek, *I. subcordata* Reissek, *I. theezans* Mart. and *I. integrifolia* Hort. being well supported). The species *I. dumosa* Reissek, *I. conocarpa* Reissek, *I. glabra* A.Gray, *I. argentina*, *I. affinis* Gardn., *I. rivularis* Gardn., *I. trichothyrsa* Loes. and *I. amara* (Vell.) Loes. form another group, but without support. *Ilex paraguariensis* A.St.-Hil. and *I. paraguariensis* var. *vestita* Loes. were the only two South American taxa that did not group with this clade, and joined with some North and Central American species (American clade), as in the MP analyses, with posterior probability (PP) 96.

The American clade is composed of a small clade formed by *I. repanda* Griseb., *I. nitida* (Vahl) Maxim., *I. discolor* Hemsl. and *I. tolucana* Hemsl. from Central America (BS 86, PP 100), and a well supported clade (BS 89, PP 100) formed by the North American taxa *I. opaca* Aiton, *I. myrtifolia* Walter, *I.* × *attenuata* Ashe and *I. cumulicola* Small, the Central American species *I. dugesii* Fernald and *I. rubra* S.Watson, and *I. paraguariensis* and *I. paraguariensis* var. *vestita*.

The European and some Asian species and *I. perado* (from the Canary Islands), form a large clade (Euroasiatic Clade) with weak support (PP 75). Two groups composed of deciduous species are present: one formed by the North American species *I. longipes*, *I. decidua* and *I. montana*, with *I. collina* being the sister taxon, with moderate bootstrap support (BS 81), and the other with strong support (PP 100) composed of the Asian species *I. fragilis*, *I. serrata*, *I. geniculata* and *I. nipponica* and the North American species *I. verticillata*. Some Asian and American species form small groups in the strict consensus tree (not shown). *Ilex purpurea*, *I. glabra*, *I. pedunculosa*, *I. canariensis*, *I. spicata*, *I. zygophylla*, *I. shennongjiaensis*, *I. cissoidea*, *I. tsoii*, *I. revoluta*, *I. aculeolata* and *I. coriacea* were unresolved.

In the combined *psbA-trnH* and *atpB-rbcL* data analysis, a total of 1456 characters was analysed,

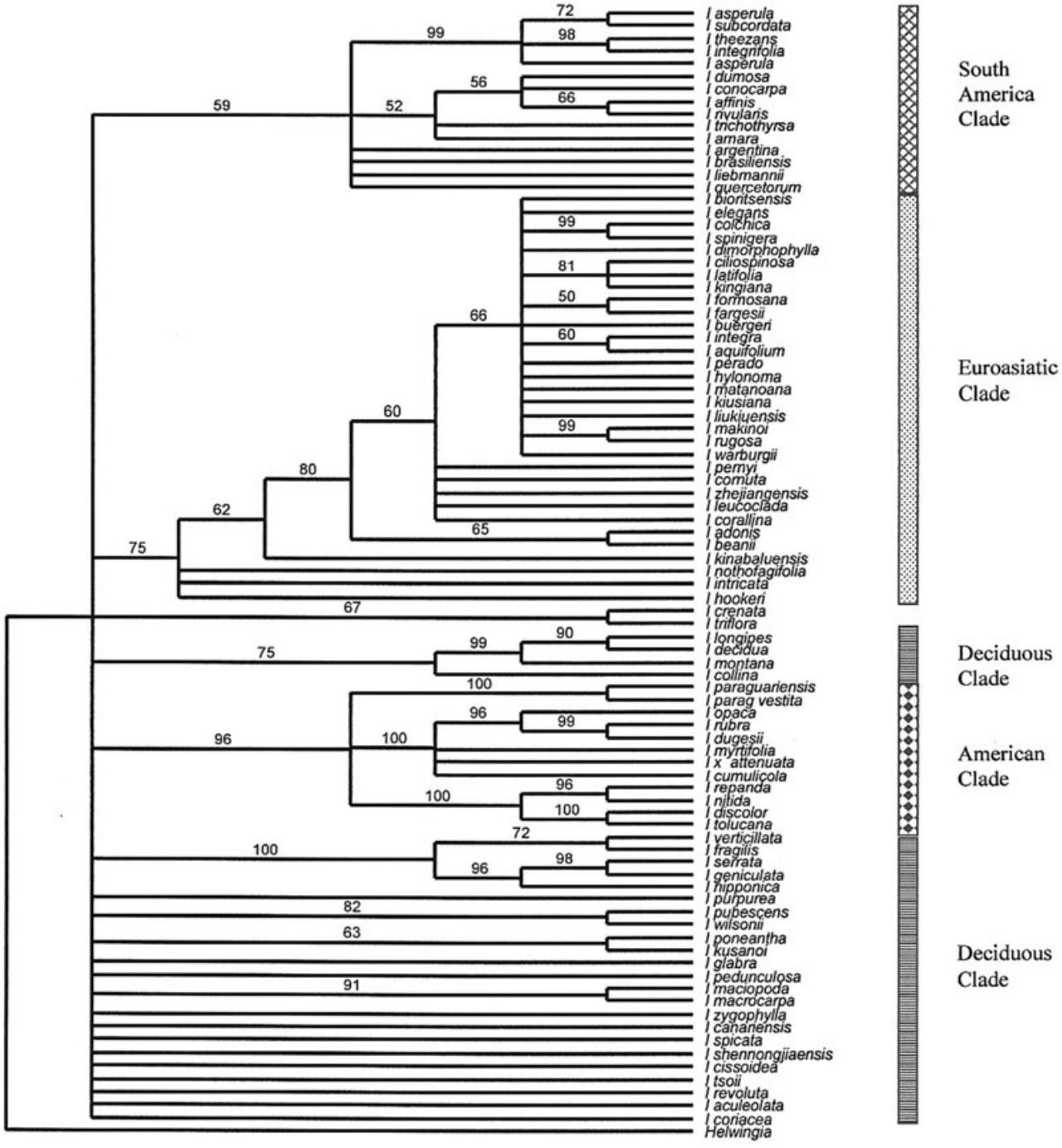

Figure 1. Majority rule consensus tree from the Bayesian analysis of the *psbA-trnH* data.

215 of which were variable and 110 were potentially parsimony informative. The parsimony search was limited to 10 000 replications. It generated 52 910 equally parsimonious trees, with a length of 519, CI of 0.71, CI excluding uninformative characters of 0.47 and RI of 0.76. The strict consensus tree of the 52 910 parsimonious trees is shown in Figure 2.

The parsimonious consensus tree calculated from the combined data analyses is more resolved than the parsimonious consensus tree based only on *psbA-trnH* data. The main difference concerns the positions of the South American species, except for *I. paraguariensis*, in a clade with some North and Central American species (as observed in Fig. 1) with a support of BS 75. The Asian and other North

American species are placed in the internal groups on the topology of the tree, most with high bootstrap support. *Ilex perado* is placed in a Euroasiatic clade (BS 76), whereas *I. canariensis* shows a basal position. The group composed of *I. longipes*, *I. decidua* and *I. montana* (also present in Fig. 1) is well supported by a high bootstrap value (BS 97). *Ilex collina* forms a sister group to the latter, but this group is not well supported (BS 61). The Asian species *I. fragilis*, *I. serrata*, *I. geniculata* and *I. nipponica* and North American species *I. verticillata* form a weakly supported clade (BS 56), with *I. zygophylla* as the sister species. Another weakly supported clade (BS 58) is composed of the Asian species *I. pubescens*, *I. wilsonii*, *I. macrocarpa*, *I. aculeolata*, *I. spicata* and *I. tsoii*.

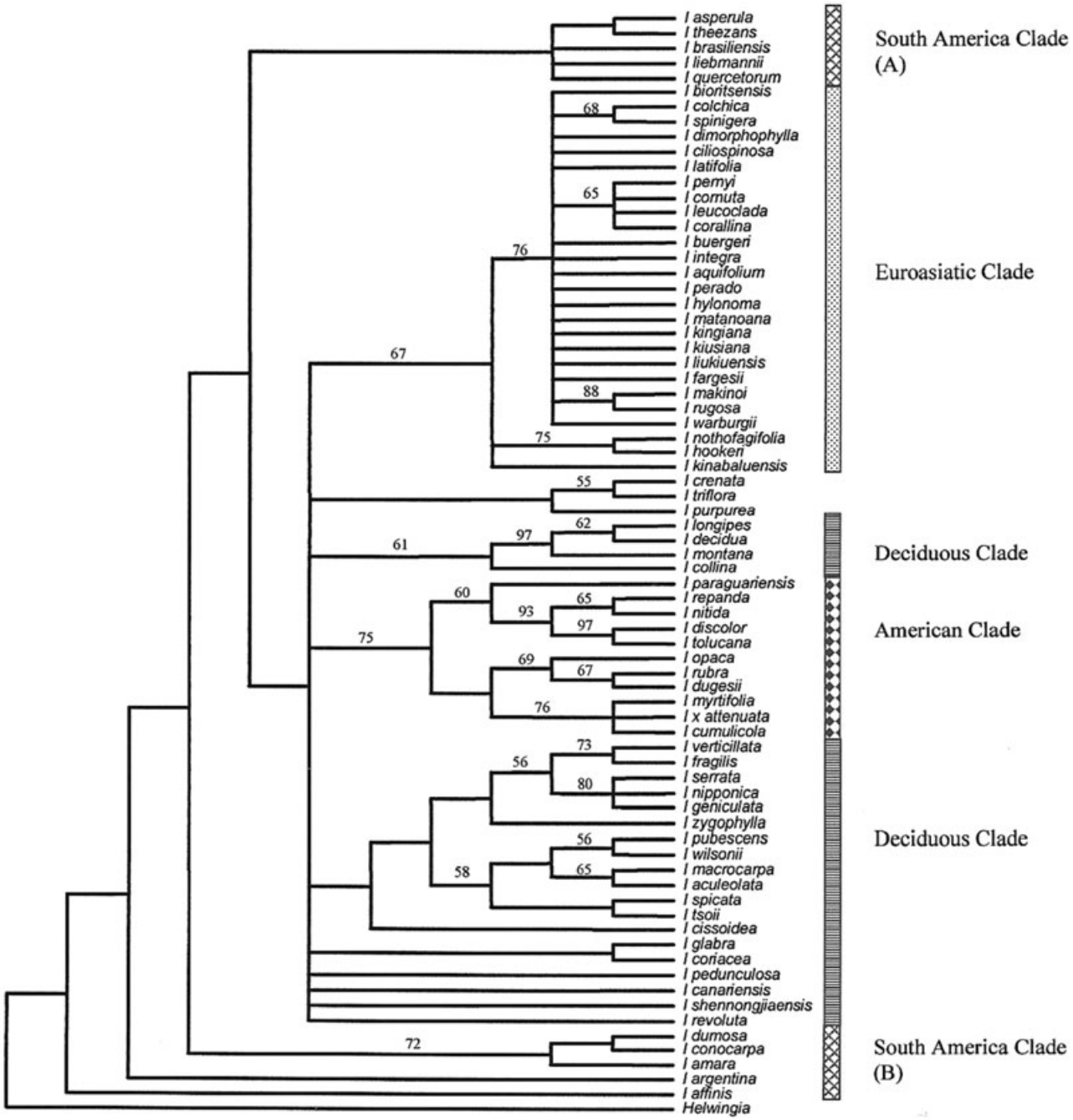

Figure 2. Strict consensus tree of the 52 910 equally parsimonious trees for *psbA-trnH* and *atpB-rbcL* combined. Bootstrap values are above the branches. Tree length, 519 steps.

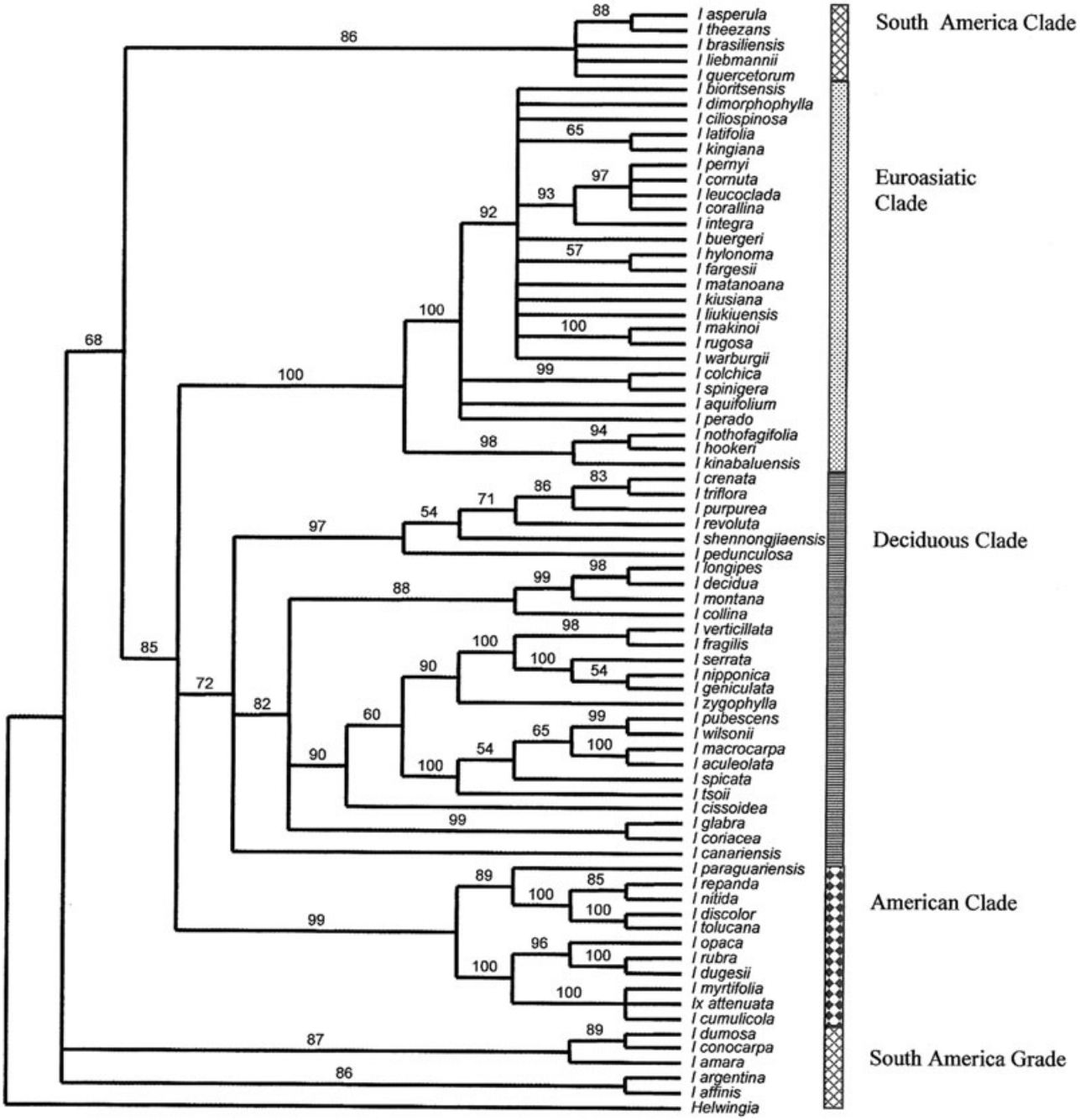

Figure 3. Majority rule consensus tree from the Bayesian analysis of the *psbA-trnH* and *atpB-rbcL* data combined.

The Bayesian analysis resulting from the combined data (*psbA-trnH* and *atpB-rbcL* including gaps) was performed with partitioning data sets with the respective models. The selected model for *atpB-rbcL* data (–ln L = 1985.8889) was F81 plus a proportion of invariable sites (I) and a gamma shape parameter (I = 0.6559; gamma = 0.9213). Stationary conditions were reached around generation 73 700, with the first 737 trees being discarded. The majority rule consensus tree calculated from Bayesian analysis is shown in Figure 3. The genus *Ilex* forms four major distinct groups. The first is composed of South and Central American species *I. asperula*, *I. theezans*, *I. brasiliensis*, *I. liebmannii* and *I. quercetorum* (South American clade A). This group is the sister group of a clade (PP 85) composed of three smaller clades. The first is composed of European and some Asian species plus *I. perado* (from the Canary Islands), forming a main

clade (Euroasiatic clade) with high support (PP 100); the second (PP 99) is formed by *I. paraguariensis* plus North and Central American species, constituting the American clade (also present in *psbA-trnH* analysis); the third group (deciduous clade) is formed by Asian and North American species plus *I. revoluta* and *I. zygophylla* (from Borneo) and *I. canariensis* (from the Canary Islands), showing a plesiomorphic insertion which ranges from 53 bp in *I. glabra* to 74 bp in *I. serrata*. This group contains all deciduous species included in the analysis. This indel was present only in some North American and Asian species, *I. canariensis* plus *H. japonica*. The South American species *I. dumosa, I. conocarpa* Reissek and *I. amara* Loes. form a small clade with *I. argentina* Lillo and *I. affinis*, and this South American group was in the first branching position.

DISCUSSION

Phylogenetic signal of *psbA-trnH* data

The plastid *psbA-trnH* intergenic spacer has been shown to evolve at high evolutionary rates and provides phylogenetic information for many groups of plants, including Magnoliaceae (Azuma *et al.*, 2001), Mimosoideae (Miller *et al.*, 2003), Penaeaceae and related families (Schönenberger & Conti, 2003) and Ericaceae (Brown *et al.*, 2006). In Lecythidaceae, the *psbA-trnH* region was more divergent than other plastid intergenic spacers, such as *trnS-trnG*, *psbB-psbH*, *atpB-rbcL*, *trnL-trnF* and *5′rpS12-rpL20* (Hamilton *et al.*, 2003). Despite this, the *psbA-trnH* region showed a slow evolutionary rate in *Ilex*. Considering the small number of potentially parsimonious characters, the search for the most parsimonious trees was limited to 10 000 replications, generating a large number of equally parsimonious trees. The strict consensus tree from the parsimonious analysis was weakly resolved and this discussion focuses on the Bayesian analysis, which produced a more resolved tree from the *psbA-trnH* data.

Previous studies using *atpB-rbcL* (Cuénoud *et al.*, 2000) and *rbcL* (Manen, Cuénoud & Martinez, 1998) to investigate the genus *Ilex* also detected a low rate of nucleotide divergence for the genus. This could be explained by the extinction of early branching lineages demonstrated in a relative rate test of nucleotide substitution (Cuénoud *et al.*, 2000), by bottleneck effects and by the reduction of effective population size (seeds are dispersed by birds at longer distances) by many species occurring in small populations (Martin, 1977). In addition, possible lineage sorting and interlineage hybridization events have occurred (Manen *et al.*, 2002). These factors impose obstacles to the phylogenetic study of this genus, as the knowledge about the relationships among the species shows a high diversity of morphological structures and a complex taxonomy. Some evolutionary relationships among species of the genus *Ilex* remain unknown.

Several indels have been observed in the *psbA-trnH* intergenic spacer, and this appears to be a common characteristic of this region (Aldrich *et al.*, 1988; Mast & Givnish, 2002; Miller *et al.*, 2003; Kyndt *et al.*, 2005; Shaw *et al.*, 2005; Winkworth & Donoghue, 2005). Indels in *psbA-trnH* are sometimes ambiguous or may generate a relatively large amount of homoplasy as a result of apparent indel 'hot spots' with numerous repeating and overlapping indels (Shaw *et al.*, 2005). In *Ilex*, one such indel indicates a divergent trend of Asian and North American species plus *I. canariensis* (the deciduous clade; Fig. 3). The size of this gap varies considerably in the deciduous clade, but its detection in only these species indicates that there are different Asian lineages giving rise to European and to North American lineages.

Phylogenetic relationships of Ilex species

Despite the high level of conservation of the *psbA-trnH* region in *Ilex* species and the consequent difficulties for phylogenetic studies using this DNA sequence, *psbA-trnH* possesses some important characters that should be exploited in this type of study. This is the case for an indel event corresponding to an insertion present only in deciduous and related species, some Asian species and *I. canariensis*. This insertion does not occur in South/Central American species, or their North American relatives or in the Eurasian Clade (Figs 2, 3). This can be explained by the existence of different lineages evolving in Asia. One lineage is associated with the European species plus *I. perado*, and another lineage is related specifically to American and deciduous species. The North American species (without this insertion) could have originated from the first Asian lineage (based on the lack of the insertion) or from South/Central American species. This last hypothesis is corroborated by a well-supported clade composed of *I. paraguariensis* (from South America) and some North and Central American species.

The origin of the South American species is an important question still unanswered by phylogenetic studies of the genus *Ilex*. Different analyses have led to different phylogenetic hypotheses concerning their evolution: almost all South American species could share a common origin (Fig. 1), or could constitute the early branching taxa in *Ilex* (Figs 2, 3). Although this observation can be explained by different data for each analysis, it cannot answer the question. Previous studies have proposed that the South American species comprise two lineages: one related to American species and the other to Asian species (Manen

et al., 2002), and the different possible relationships among South American species have been discussed in other papers on *Ilex* phylogeny. In a phylogenetic study using ITS sequences from South American species (Gottlieb, Giberti & Poggio, 2005), *I. argentina*, *I. brasiliensis*, *I. brevicuspis* Reissek, *I. integerrima* Reissek, *I. microdonta* Reissek, *I. pseudobuxus* Reissek, *I. taubertiana* Loes. and *I. theezans* were related in all the different conditions under which the study was conducted. However, the position of other South American species, including *I. paraguariensis*, *I. dumosa* varieties and one sample of *I. argentina*, was highly dependent on the parameters used in the analyses. The sister group of the South American clade was only the Asian species *I. pedunculosa* in most parameter sets assayed, but with low bootstrap support values (67). The *I. dumosa* varieties were more closely related to the eastern Asian species *I. crenata* and *I. mutchagara* Makino. Despite the presence of different ITS sequences in some *Ilex* species, the analyses performed using amplified fragment length polymorphism (AFLP) markers (Gottlieb *et al.*, 2005) showed that the samples from the same morphospecies formed distinct clades. In such cases, the distinction between *I. dumosa* var. *dumosa* Reissek and *I. dumosa* var. *guaranina* Loes. was not detected.

The data presented here confirm the close relationship between *I. theezans* and *I. brasiliensis* (Figs 2, 3), as pointed out by Giberti (1998a) and Gottlieb *et al.* (2005). After examining the type material cited by Coelho & Mariath (1996), it is possible to observe a close proximity between *I. dumosa* and *I. amara* (possibly belonging to the same morphospecies), also hypothesized by Andrews (1985) and Giberti (1998b). Our results are in agreement with this conclusion (Figs 2, 3), in spite of the inclusion of *I. conocarpa* in the clade that groups those species.

The European species of *Ilex* are closely related to the Asian species, and it is possible to assign several ancestral taxa to them. *Ilex spinigera* Loes. and *I. colchica* Pojark. fell together in all analyses performed, and that they could share the same ancestor, different from European species: *I. aquifolium*. The Eurasian clade described by Cuénoud *et al.* (2000) was present in *psbA-trnH* Bayesian analysis and in the combined *psbA-trnH* and *atpB-rbcL* data analyses.

The deciduous clade described by Cuénoud *et al.* (2000) was not evident in the analysis using *psbA-trnH*. In the present study, the deciduous species only formed two clades. In the Bayesian analysis using the combined data, the deciduous clade includes all deciduous species and some evergreen species.

The evolution of the Canary Island species included in this study should be more fully investigated, but they probably have different origins. *Ilex canariensis* seems to be related to one group of East Asian and North American species. All species of this group plus *I. canariensis* have an insertion in the *psbA-trnH* intergenic spacer, but the origin of this species is still uncertain. The second species, *I. perado*, is related to another Asian group that probably originated from European species and does not share the same insertion present in *I. canariensis*. In a similar case, two Macaronesian endemic species of the genus *Lavatera* L. (Malvaceae), *L. phoenicea* Vent. and *L. acerifolia* Cav., represent two independent introductions into the Canary Islands (Fuertes-Aguilar *et al.*, 2002). Incongruence between plastid and nuclear phylogenetic analyses suggests that hybridization may have played a role in the evolution of *L. acerifolia*, like the Macaronesian *Ilex* species (Fuertes-Aguilar *et al.*, 2002). Direct and indirect evidence indicates that hybridization among related taxa is rather common, and has played an important role in the evolutionary history of Macaronesian plants (Herben, Suda & Munclinger, 2005).

Taxonomic implications

This work and other previous molecular studies concerning the genus *Ilex* (Cuénoud *et al.*, 2000; Manen *et al.*, 2002; Gottlieb *et al.*, 2005) do not agree with the classifications of *Ilex* proposed by Loesener (1901, 1908, 1942) or Galle (1997).

The phylogenetic significance and taxonomic value of some flavones and flavonols in *Ilex* species have been studied (Martinez, Pelotto & Basualdo, 1997). Flavonols were widely distributed in the studied species of the genus *Ilex*, but flavones were present in only *I. belizensis* Lundell and *I. leucoclada* Makino. Aglycones showed a tendency to accumulate differentially in each region. Isorhamnetin was present at a higher frequency in American species than in the Asiatic group. Kaempferol was less frequent in species from Central America, but was found with quercetin and without isorhamnetin in *I. mitis* Radlk. from Africa. Because of their sporadic occurrence, these flavones could be interesting markers to determine the degree of association among allied species. The presence of apigenin and luteolin in *I. belizensis* carries potential taxonomic value in identifying the limits of the alliance in a group in which misidentification is frequent and the possibility of finding new species is high. A chemotaxonomic analysis proved that it is possible to discriminate South American *Ilex* species by multivariate analysis of their metabolite fingerprints (Choi *et al.*, 2005). The major metabolites contributing to this discrimination are arbutin, caffeine, phenylpropanoids and theobromine. Arbutin has been found to be a biomarker for *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, *I. microdonta*, *I. pseudobuxus*, *I. taubertiana* and *I. theezans*. However, arbutin was not present in detect-

able amounts in samples of *I. dumosa* var. *dumosa*, *I. dumosa* var. *guaranina* and *I. paraguariensis* var. *paraguariensis*. In addition to these major phenolic metabolites, a number of putative minor metabolites also play a role in differentiating between *Ilex* species. These studies and other papers that use combined molecular and morphological data are necessary to contribute to a taxonomic review of the genus *Ilex*, as the current system is not corroborated by molecular phylogenetic studies.

CONCLUSIONS

A high level of conservation was observed in the plastid *psbA-trnH* intergenic spacer in *Ilex* species, despite its high rate of evolution and potential phylogenetic informative value in other groups of plants (Azuma *et al.*, 2001; Holderegger & Abbott, 2003; Schönenberger & Conti, 2003; Kim *et al.*, 2004; Butterworth & Wallace, 2005; Bruyns *et al.*, 2006). However, this fragment included an informative (and variable) insertion in deciduous species and their relatives. Some Asian species plus *I. canariensis*, South/Central American species and North American relatives and the Eurasian clade do not share this indel. This suggests the existence of different lineages of species evolving in Asia. The Canary Island species have different origins, being related to different Asian lineages. European species seem to be related to one of the Asian lineages. The origin of South American species still remains unknown, but there are different lineages in this region.

ACKNOWLEDGEMENTS

We thank the Porto Alegre Botanical Garden, the Royal Botanic Gardens, Kew, Dr Ching-Yeh Hu and Dr Armando Cervi for providing plant samples, Ing. L.D. Belingheri (INTA, Cerro Azul, Argentina) for providing several *Ilex* specimens and Dr José Artur Bogo Chies (UFRGS), Dr Giancarlo Pasquali (UFRGS) and Dr Maurício Reis Bogo (PUC-RS) for the sequencing facilities. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS), Secretaria da Ciência e Tecnologia do Estado do Rio Grande do Sul (SCT-RS) and Instituto Brasileiro do Meio Ambiente (IBAMA)/Ilópolis.

REFERENCES

Aldrich J, Cherney BW, Merlin E, Christopherson L. 1988. The role of insertions/deletions in the evolution of the intergenic region between *psbA* and *trnH* in the chloroplast genome. *Current Genetics* **14:** 137–146.

Andrews S. 1985. A checklist of the Aquifoliaceae of Bahia. *Rodriguésia* **37:** 34–44.

Azuma H, García-Franco JG, Rico-Gray V, Thien LB. 2001. Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. *American Journal of Botany* **88:** 2275–2285.

Barral G, Poggio L, Giberti GC. 1995. Chromosome numbers and DNA content from *Ilex argentina* (Aquifoliaceae). *Boletín de la Sociedad Argentina de Botánica* **30:** 243–248.

Brown GK, Craven LA, Udovicic F, Ladiges PY. 2006. Phylogeny of *Rhododendron* section *Vireya* (Ericaceae) based on two non-coding regions of cpDNA. *Plant Systematics and Evolution* **257:** 57–93.

Bruyns PV, Mapaya RJ, Hedderson T. 2006. A new subgeneric classification for *Euphorbia* (Euphorbiaceae) in southern Africa based on ITS and *psbA-trnH* sequence data. *Taxon* **55:** 397–420.

Butterworth CA, Wallace RS. 2005. Molecular phylogenetics of the leafy cactus genus *Pereskia* (Cactaceae). *Systematic Botany* **30:** 800–808.

Choi YH, Sertic S, Kim H, Wilson EG, Michopoulos F, Lefeber AWM, Erkelens C, Kricun SDP, Verpoorte R. 2005. Classification of *Ilex* species based on metabolomic fingerprinting using nuclear magnetic resonance and multivariate data analysis. *Journal of Agricultural and Food Chemistry* **53:** 1237–1245.

Coelho GC, Mariath JEA. 1996. Inflorescence morphology of *Ilex* L. (Aquifoliaceae) species from Rio Grande do Sul, Brazil. *Feddes Repertorium* **107:** 19–30.

Cuénoud P, Martinez MADP, Loizeau PA, Spichiger R, Andrews S, Manen J-F. 2000. Molecular phylogeny and biogeography of the genus *Ilex* L. (Aquifoliaceae). *Annals of Botany* **85:** 111–122.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19:** 11–15.

Fuertes-Aguilar J, Ray MF, Francisco-Ortega J, Santos-Guerra A, Jansen RK. 2002. Molecular evidence from chloroplast and nuclear markers for multiple colonizations of *Lavatera* (Malvaceae) in the Canary Islands. *Systematic Botany* **27:** 74–83.

Galle FC. 1997. *Hollies: the genus Ilex*. Portland, OR: Timber Press.

Giberti GC. 1998a. Hallazgo de *Ilex brasiliensis* (Aquifoliaceae) en la Argentina. *Boletín de la Sociedad Argentina de Botánica* **33:** 137–140.

Giberti GC. 1998b. Aquifoliaceae paraguayas: notas críticas en el género *Ilex* L. *Candollea* **43:** 417–420.

Gottlieb AM, Giberti GC, Poggio L. 2005. Molecular analyses of the genus *Ilex* (Aquifoliaceae) in southern South America, evidence from AFLP and ITS sequence data. *American Journal of Botany* **92:** 352–369.

Hamilton MB, Braverman JM, Soria-Hernanz DF. 2003. Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. *Molecular Biology and Evolution* **20:** 1710–1721.

Herben T, Suda J, Munclinger P. 2005. The ghost of hybridization past: niche pre-emption is not the only explanation of apparent monophyly in island endemics. *Journal of Ecology* **93:** 572–575.

Holderegger R, Abbott RJ. 2003. Phylogeography of the Arctic-Alpine *Saxifraga oppositifolia* (Saxifragaceae) and some related taxa based on cpDNA and its sequence variation. *American Journal of Botany* **90:** 931–936.

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* **17:** 754–755.

Judd WS, Campbell CS, Kellogg EA, Stevens PF. 1999. *Plant systematics – a phylogenetic approach*. Sunderland, MA: Sinauer Associates, Inc.

Kim SC, Lu CT, Lepschi BJ. 2004. Phylogenetic positions of *Actites megalocarpa* and *Sonchus hydrophilus* (Sonchinae: Asteraceae) based on ITS and chloroplast non-coding DNA sequences. *Australian Systematic Botany* **17:** 73–81.

Klak C, Khunou A, Reeves G, Hedderson T. 2003. A phylogenetic hypothesis for the Aizoaceae (Caryophyllales) based on four plastid DNA regions. *American Journal of Botany* **90:** 1433–1445.

Kyndt T, Droogenbroeck BV, Romeijn-Peeters E, Romero-Motochi JP, Scheldeman X, Goetghebeur P, Damme PV, Gheysen G. 2005. Molecular phylogeny and evolution of Caricaceae based on rDNA internal transcribed spacers and chloroplast sequence data. *Molecular Phylogenetics and Evolution* **37:** 442–459.

Loesener T. 1901. Monographia aquifoliacearum. *Nova Acta Academiae Caesareae Leopoldina – Carolinae germanicae Naturae Curiosorum* **78:** 1–598.

Loesener T. 1908. Monographia aquifoliacearum. *Nova Acta Academiae Caesareae Leopoldina – Carolinae germanicae Naturae Curiosorum* **89:** 5–314.

Loesener T. 1942. Aquifoliaceae. In: Engler A, Prantl K, eds. *Die Natürlichen Pflanzenfamilien* 2nd ed. **20b:** 36–86. Leipzig: Wilhelm Engelmann.

Loizeau PA. 1994. Les Aquifoliaceae péruviennes (eléments pour une révision des Aquifoliaceae néotropicales). *Boissiera* **48:** 5–306.

Loizeau PA, Spichiger R. 1992. Proposition d'une classification des inflorescences d'*Ilex* L. (Aquifoliaceae). *Candollea* **47:** 97–112.

Manen JF. 2004. Are both sympatric species *Ilex perado* and *Ilex canariensis* secretly hybridizing? Indication from nuclear markers collected in Tenerife. *BMC Evolutionary Biology* **4:** 46–58.

Manen J-F, Boulter MC, Naciri-Graven Y. 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 Systematics and Evolution* **235:** 79–98.

Manen JF, Cuénoud P, Martinez MDR. 1998. Intralineage variation in the pattern of *rbcL* nucleotide substitution. *Plant Systematics and Evolution* **211:** 103–112.

Martin HA. 1977. The history of *Ilex* (Aquifoliaceae) with special reference to Australia: evidence from pollen. *Australian Journal of Botany* **25:** 655–673.

Martinez MADP, Pelotto JP, Basualdo N. 1997. Distribution of flavonoid aglycones in *Ilex* species (Aquifoliaceae). *Biochemical Systematics and Ecology* **25:** 619–622.

Mast AR, Givnish TJ. 2002. Historical biogeography and the origin of stomatal distributions in *Banksia* and *Dryandra* (Proteaceae) based on their cpDNA phylogeny. *American Journal of Botany* **89:** 1311–1323.

Miller JT, Grimes JW, Murphy DJ, Bayer RJ, Ladiges PY. 2003. A phylogenetic analysis of the Acacieae and Ingeae (Mimosoideae: Fabaceae) based on *trnK*, *matK*, *psbA-trnH*, and *trnL/trnF* sequence data. *Systematic Botany* **28:** 558–566.

Nylander JAA. 2004. *MrModeltest v2. Program distributed by the author*. Uppsala: Evolutionary Biology Centre, Uppsala University.

Ronquist F, Huelsenbeck JP, Mark PVD. 2005. *MrBayes 3.1 manual*. Available at: http://mrbayes.csit.fsu.edu/mb3.1_manual.pdf

Saghai-Maroof MF, Soliman KM, Jorgensen RA, Allard RW. 1984. Ribosomal DNA spacer-length polymorphisms in barley. Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **81:** 8014–8018.

Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84:** 1120–1136.

Schönenberger J, Conti E. 2003. Molecular phylogeny and floral evolution of Penaeaceae, Oliniaceae, Rhynchocalycaceae and Alzateaceae. *American Journal of Botany* **90:** 293–309.

Setoguchi H, Watanabe I. 2000. Intersectional gene flow between insular endemics of *Ilex* (Aquifoliaceae) on the Bonin Islands and the Ryukyu Islands. *American Journal of Botany* **87:** 793–810.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92:** 142–166.

Smissen RD, Breitwieser I, Ward JM. 2004. Phylogenetic implications of trans-specific chloroplast DNA sequence polymorphism in New Zealand Gnaphalieae (Asteraceae). *Plant Systematics and Evolution* **249:** 37–53.

Swofford DL. 2002. *PAUP: phylogenetic analysis using parsimony*. Version 4.0b. Champaign, IL: Illinois Natural Survey.

Thompson JD, Higgins DG, Gibson TJ. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22:** 4673–4680.

Winkworth RC, Donoghue MJ. 2005. *Viburnum* phylogeny based on combined molecular data: implications for taxonomy and biogeography. *American Journal of Botany* **92:** 653–666.