

Molecular Analyses of the Genus *Ilex* (Aquifoliaceae) in Southern South America, Evidence from AFLP and ITS Sequence Data

Author(s): Alexandra M. Gottlieb, Gustavo C. Giberti and Lidia Poggio

Source: *American Journal of Botany*, Vol. 92, No. 2 (Feb., 2005), pp. 352-369

Published by: Botanical Society of America, Inc.

Stable URL: <https://www.jstor.org/stable/4123880>

Accessed: 18-12-2018 17:13 UTC

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/4123880?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Botanical Society of America, Inc. is collaborating with JSTOR to digitize, preserve and extend access to *American Journal of Botany*

MOLECULAR ANALYSES OF THE GENUS *ILEX* (AQUIFOLIACEAE) IN SOUTHERN SOUTH AMERICA, EVIDENCE FROM AFLP AND ITS SEQUENCE DATA¹

ALEXANDRA M. GOTTLIEB,^{1,3} GUSTAVO C. GIBERTI,² AND
LIDIA POGGIO¹

¹Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Ciudad Universitaria, Pabellón II, 4to piso, Departamento de Ecología, Genética y Evolución, Lab. 62, C1428EHA, Buenos Aires, Argentina; and ²Instituto de Química y Metabolismo del Fármaco (IQUIMEFA), Junín 956, 1113-Buenos Aires, Argentina

In order to clarify the relationships among southern South American (sSA) representatives of the genus *Ilex*, an amplified fragment length polymorphism (AFLP) analysis was accomplished. In addition, the phylogenetic relationships of the species were studied using ribosomal internal transcribed spacer (ITS) sequence data alone and in combination with AFLP data, taking into account the possible existence of paralogous sequences and the influence of alignment parameters. To explore stability of phylogenetic hypotheses, a sensitivity analysis was performed using 15 indel-substitution models. Within each species assayed, the AFLPs allowed the recognition of several diagnostic bands. Furthermore, the AFLP analysis revealed that individuals belonging to the same morpho-species formed coherent clades. In addition, some cases of geographical association were noted. Studies on ITS sequences revealed divergence between data obtained herein and sequence data downloaded from GenBank. The sensitivity analyses yielded different interspecific hypotheses of relationships. Notwithstanding, analyses of the ITS data alone and in combination with AFLPs, rendered clades stable to variation in the analytical parameters. Topologies obtained for the AFLPs, the ITS data alone and the combined analyses, demonstrated the existence of a group formed by *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, and *I. theezans*, and that *I. dumosa* and *I. paraguariensis* were distantly related to the former. Incongruence with traditional taxonomical treatments was found.

Key words: AFLP; direct optimization; *Ilex*; ITS; sensitivity analysis.

The plant genus *Ilex* L., which belongs to the holly family, Aquifoliaceae Bartl., comprises deciduous and evergreen bushes or trees with economic importance as crops and ornamentals. These plants are functionally dioecious; rudimentary pistils in staminate flowers and sterile stamens in pistillate flowers are borne on different plants. A remarkable morphological diversity has been described in the genus, including trees to creeping prostrate plants or epiphytic shrubs. It has more than 400 species, dispersed mainly in America and Eurasia, but also in Oceania and Africa (Giberti, 1994b).

South America is considered one of the main areas of diversification of *Ilex*, together with East Asia (Loesener, 1942; Cuénoud et al., 2000). In southern South America (sSA) the species are mostly found in areas with tropical or subtropical climates, mainly in northeastern Argentina, southeastern Brazil and eastern Paraguay. Twelve species have been described for sSA (Loesener, 1901; Lillo, 1911; Edwin and Reitz, 1967; Giberti, 1998), namely, *Ilex affinis* Gardner, *I. argentina* Lillo, *I. brasiliensis* (Sprengel) Loes., *I. brevicuspis* Reissek, *I. chamaedryfolia* Reissek, *I. dumosa* Reissek, *I. integerrima* (Vell.)

Reissek, *I. microdonta* Reissek, *I. paraguariensis* A. St. Hil., *I. pseudobuxus* Reissek, *I. taubertiana* Loes., and *I. theezans* C. Martius ex Reissek. Of those mentioned, *I. paraguariensis* is the most relevant from an economic perspective. In Argentina, southern Brazil, Paraguay, and Uruguay, the aerial parts of this plant are used to prepare a very popular beverage called mate, which is highly appreciated for its flavor and stimulating properties due to its caffeine and theobromine contents (Filip et al., 2001).

About a hundred years after Linnaeus (1753) published the diagnosis for the genus *Ilex*, Gray (1856) attempted to establish an infrageneric classification. In 1861, Reissek published the greatest species inventory for the time, followed by Hooker (1862) who reported the occurrence of about 145 species worldwide. Loesener (1891, 1901, 1908, 1942) published a series of seminal studies recording almost 280 species of *Ilex*, dispersed globally, and recognizing several subgeneric ranks although under an unconventional hierarchical arrangement. Since then, several authors have recognized some inaccuracies in Loesener's system and proposed emendations in the systematics of the genus (Rehder, 1927; Hu, 1949, 1950; Krüssmann, 1962; Baas, 1975; Lobreau-Callen, 1975; Martin, 1977). Species delimitation in *Ilex*, based on overall morphological similarity, is complicated due to the various concepts used by different authors to define the taxa and to the lack of complete information, particularly for evergreen species (Galle, 1997).

In order to investigate the relationships among the species of *Ilex* found in sSA, we implemented the amplified fragment length polymorphism (AFLP) technique (Vos et al., 1995). In addition, we studied the phylogenetic relationships of the species using ribosomal internal transcribed spacer (ITS) sequence data alone and in combination with AFLP data, taking into account the possible existence of paralogous sequences and the influence of alignment parameters.

¹ Manuscript received 11 December 2003; revision accepted 17 September 2004.

The authors are indebted to L. D. Belingheri, S. D. Prat Kricun, and R. M. Mayol from EE. INTA Cerro Azul (Misiones, Argentina) who kindly provided specimens from living collections; to N. Juvé and P. Rubio of University of Alcalá (Madrid, Spain) who supported the sequencing, to M. E. Rolón from the "Establecimiento Las Marias" (Corrientes, Argentina) who facilitated commercial lines of "mate," to S. A. Catalano and V. A. Confalonieri for their invaluable suggestions. We also want to thank three anonymous reviewers whose suggestions helped to improve this manuscript. This work was supported from grants of the "Agencia Nacional de Investigaciones Científicas y Técnicas" (Argentina) (PICT 01-4443) and from CONICET (PIP 0559/98).

³ Author for correspondence. E-mail: gottlieb@bg.fcen.uba.ar.

Beside the general use of AFLP markers in the estimation of genetic diversity and differentiation among individuals, populations, and plant species, they have also been successfully used in phylogenetic analysis between closely related species (Cervera et al., 1998; Kardolus et al., 1998; Mace et al., 1999; Negi et al., 2000; Ren and Timko, 2000; Koopman et al., 2001). Because of the rapidity with which reliable high resolution markers can be generated, the AFLP technique is considered a powerful tool for molecular studies (Mueller and Wolfenbarger, 1999). In addition, the high ratio of polymorphism generated per PCR experiment (the multiplex ratio) (Powell et al., 1996) and the random distribution of most AFLP bands across the genome (Zhu et al., 1998) are considered key features of this technique. Several drawbacks have been detected for AFLPs (Robinson and Harris, 1999), which introduce bias into the estimates, whether analyzed by phenetics or cladistics.

A fundamental requirement in phylogenetic studies based on nucleic acid or protein sequences is that the genes compared are orthologous (originating from organismal cladogenesis) (Alvarez and Wendel, 2003). Plant rDNA consists of thousands of copies or paralogues (derived from gene duplication events) spanning several loci, which evolve through gene conversion, unequal crossing over, and perhaps repeat amplification (Baldwin et al., 1995). ITS sequences have proven useful in species-level phylogenetic studies of a wide range of taxa (Alice and Campbell, 1999; Downie et al., 2000; Lia et al., 2001, among others). Within-individual rDNA polymorphisms may occur when concerted evolution is not fast enough to homogenize repeats in the face of high rates of mutation and/or recent interspecific hybridization (Campbell et al., 1997). Intraspecific variation concerning ITS sequences has been detected in several plant groups (Baldwin et al., 1995 and references therein; Buckler and Holtsford, 1996a, b; Buckler et al., 1997; Campbell et al., 1997; Hartmann et al., 2001; Mayol and Roselló, 2001). Active ITS regions have functional constraints that can help discriminate between functional and nonfunctional paralogous (i.e., pseudogene) sequences (Buckler et al., 1997). For instance, nonfunctional ITS paralogues could accumulate random substitutions, which will destabilize hairpins and reduce secondary structure stability. When nonfunctional paralogous ITS sequences are unknowingly included in the phylogenetic analyses (to the exclusion of appropriate orthologous comparisons), the resulting gene tree may confound organismal divergence events with a tracking of the history of gene duplication (Alvarez and Wendel, 2003). Erroneous assessment of orthology and paralogy will lead to phylogenetic incongruence. Hence, discrimination of inactive paralogues and pseudogenes becomes crucial.

Another critical step in molecular phylogenetic studies is the alignment of sequence data, in which primary homology statements are being established. However, it is typically performed without considering the effects of analytical parameters, such as indel and base transformation values, on the phylogenetic inference. In the case of multiple sequence alignments, different parameters may yield alternative alignments, and these primary homology hypotheses may support distinct topologies (Doyle and Davies, 1998). Indeed, Morrison and Ellis (1997) showed that different multiple sequence alignment methods were responsible for the variation encountered among resultant phylogenetic trees topologies, and this variation was more important than the one found when different methods of phylogenetic reconstruction were applied to the same align-

ments. Because parameter choice is arbitrary but unavoidable when doing algorithmic DNA sequence comparisons, its exploration becomes essential (Giribet and Wheeler, 1999; Giribet and Ribera, 2000). Herein, we used ITS sequences to study the phylogenetic relationships of sSA taxa and to explore the influence of applying different insertion/deletion (indel or gap) and substitution models (sensitivity analysis sensu Wheeler, 1995) on phylogenetic inference. Phylogenetic analyses were accomplished through direct optimization of DNA sequences (Wheeler, 1996). The sensitivity analysis is a way to explore the data and to discern between stable relationships (those supported throughout the parameters range) and unstable relationships (those appearing only under particular conditions), allowing the formulation of more robust hypotheses (Giribet and Ribera, 2000) than in customary phylogenetic analyses. Direct optimization is a method that does not disconnect the alignment step from the tree-building step and that generalizes phylogenetic character analysis to include indel events (Giribet, 2001, 2003). By doing this, indels appear as transformations (not as states) linking ancestral and descendent nucleotide sequences (Giribet and Wheeler, 1999).

MATERIALS AND METHODS

Specimens studied—The list of sSA *Ilex* taxa employed, together with their accession numbers, geographical origin and GenBank accession numbers of ITS sequences obtained in this study are presented in Appendix 1 (see Supplemental Data accompanying the online version of this article). Plant material was principally obtained from the Estación Experimental INTA Cerro Azul (EEINTACA) Germplasm Bank (Misiones, Argentina). Leaf samples were also collected from natural populations from Northern Argentina; voucher specimens of these samples were deposited at the BACP Herbarium (Centro de Estudios Farmacológicos y Botánicos, CEFyBO, Buenos Aires, Argentina). Furthermore, leaf samples of commercial lines of mate were obtained from the Establecimiento Las Marías (Corrientes, Argentina). Young leaf samples were collected from healthy plants (without evidence of fungal or insect infection) and were silica-gel preserved. Additional 42 ITS sequences were downloaded from GenBank (<http://www.ncbi.nlm.nih.gov>); for their geographical origin and accession numbers see Appendix 2 (Data Supplement accompanying the online version of this article). *Helwingia japonica* (C. P. Thunberg ex A. Murray) Dietrich and *Helwingia chinensis* Batalin were chosen as the outgroup. According to several molecular studies (Cuénoud et al., 2000; Powell et al., 2000; Savolainen et al., 2000; Albach et al., 2001; Bremer et al., 2002), *Helwingia* and *Ilex* are closely related members of Aquifoliales. Loesener's (1891, 1901, 1908, 1942) original dispositions of subgeneric ranks were used in this work (Fig. 19). The species *Ilex microdonta*, *I. pseudobuxus*, and *I. taubertiana*, were only included in the sequence analyses because accessions of these taxa were unavailable in sufficient number for AFLP analysis. Only two sSA taxa remain to be studied, namely *Ilex affinis* and *I. chamaedryfolia*, because these have not been found in the field since the 1980s (G. Giberti, Instituto de Química y Metabolismo del Fármaco, IQUM-EFA, personal communication). Attempts to obtain good quality DNA from herbarium/air-dried specimens were unsuccessful.

DNA extraction—Leaves were ground to a fine powder in liquid nitrogen and then placed in a microtube. The DNeasy Plant kit (QIAGEN Inc., Valencia, California, USA) was used for DNA extraction following the manufacturer's instructions. DNA was kept at 4°C until use and then stored at -20°C.

AFLP technique—The AFLP methodology was carried out using the AFLP Analysis System I Starter Primer kit (Gibco BRL, Life Technologies, Carlsbad, California, USA) as described in the instruction manual. Eight selective primers were combined as follows: E + AGG, M + CTT; E + ACT, M + CAC; E + AGC, M + CAT; and E + ACC, M + CAG. Other primers combinations assayed were discarded because they generated profiles in which

the amplification products were too dense to allow reliable scoring, or alternatively, generated too few amplification products. Selective amplification products were mixed with an equal volume of dye reagent (98% [v/v] formamide, 10 mM EDTA, 0.025% [w/v] bromophenol blue and 0.025% [w/v] xylene cyanol). Five microlitres were separated by electrophoresis in a Model S2 apparatus (Gibco BRL Sequencing System, Life Technologies) through 6% (w/v) polyacrylamide gels containing 5 M urea, in $1 \times$ TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8). A 30–330 bp AFLP DNA Ladder (Gibco BRL, Life Technologies) size marker was included twice or three times in each electrophoresis run. Thus, the size of AFLP bands scored ranged from 50 to 330 bp. Gels were stained with silver nitrate (Bassam et al., 1991).

ITS amplification—Amplification of ITS 1, 5.8S gene, and ITS 2 was carried out using primers ILEXFP (5'-AACAAAGGTTCCGTAGGTGA-3', Powell et al., 2000) and ITS-4 (White et al., 1990). Amplifications were performed in 50 μ L using 0.225 mM of each dNTP (Promega Corp., Madison, Wisconsin, USA), 20–25 pmol of each primer, 5 μ L of Thermophilic DNA Polymerase 10 \times Buffer (Promega Corp.), 1.5 mM MgCl₂, 1.25 units of *Taq* DNA polymerase (Promega Corp.), 1 μ L of genomic DNA template, and sterile double-distilled water. Conditions for PCR amplification were those of Vilgalys and Hester (1990). Amplifications were carried out in an Eppendorf Mastercycler (Perkin-Elmer Corp., Foster City, California, USA). Control samples without DNA template were included in each PCR run. PCR products were checked by electrophoresis of 1- μ L samples in 1% agarose gels in $1 \times$ TAE buffer (0.04 M Tris, 0.114% glacial acetic acid, 1 mM EDTA, pH 8.0). Quantification of PCR products was estimated by comparison with known amounts of a DNA molecular-size marker (Lambda *EcoRI/HindIII*, Promega Corp.) included in duplicate in all gels. Gels were stained with ethidium bromide for 15–20 min and photographed under UV light. In some cases, best amplification results were achieved by adding 6% bovine serum albumin (BSA, Promega Corp.) to the PCR reaction mix. PCR products were purified using a QIAGEN Gel Extraction kit (QIAGEN Inc.). Both strands of the complete ITS region (ITS 1–5.8S–ITS 2) were sequenced in Perkin Elmer/Applied Biosystems automatic sequencing units (ABI Prism 310 or 377) at the Molecular Biology Facility of the Universidad de Alcalá (Madrid, Spain). Electropherograms were proofread with the software BioEdit Sequence Alignment Editor (Hall, 1999). Boundaries of the coding and spacer regions were determined by comparison with published sequences of *Ilex*. In general, when two or more orthologous ITS sequences were available per sSA taxon (see Appendix 1; Supplemental Data accompanying the online version of this article), these were used to produce consensus sequences in BioEdit (Hall, 1999) to speed-up computational times of phylogenetic analyses.

Data analyses—**AFLP**—Air-dried gels were scanned and banding patterns were registered with the Gel-Pro 4 Analyzer Program (Media Cybernetics, Maryland, USA). Each AFLP band, regardless of its relative intensity, was considered as a dominant allele at a unique locus. In some cases, fragments were scored as missing data because character states could not be interpreted unambiguously. The data were extracted as a table as either present (1) or absent (0). Monomorphic bands (bands present in all individuals of a species) and diagnostic bands (monomorphic bands exclusively present in one species) were discriminated within each species and across the entire data set. Box plots of mean values; SE and SD were calculated and plotted with the program STATISTICA '99 edition (Kernel release 5.5 A, StatSoft, 1999). For these calculations, accessions with missing data in one or more primer combinations were excluded. Pairwise genetic distance (D) between all individuals was estimated according to the complementary value of Nei and Li's (1979) similarity coefficient, implemented in PAUP* (Swofford, 1998): $D = 1 - (2n_{ij} / [n_i + n_j])$, where n_{ij} is the number of shared fragments between two individuals i and j , and n_i and n_j are the total number of fragments in individual i and j , respectively. From these values, mean intra- and interspecific distance values were obtained. Furthermore, the AFLP data set for 120 accessions belonging to eight taxa was analyzed using maximum parsimony in PAUP*. Initial trees in the heuristic search by tree-bisection-reconnection (TBR) branch-swapping algorithm were built stepwise with 500 random addition

sequence replicates. Characters were equally weighted and unordered, and all other settings were defaults. The reliability of the clades was evaluated by bootstrap analysis (Felsenstein, 1985) in PAUP* with 1000 replications and all other settings as described earlier. Trees were inspected with the program TreeView (Page, 1998). The unrooted resulting strict consensus tree was rooted using *I. paraguariensis* to be consistent with ITS trees.

ITS sequence studies—We characterized and identified ITS paralogues in sSA *Ilex* species by examining their G + C content, secondary structure stabilities, and nucleotide substitution patterns and using the tree-based approach advocated by Bailey et al. (2003). To discriminate the presence of presumed active sequences from paralogues and putative pseudogenes, we first followed the procedure proposed by Mayol and Roselló (2001). Therefore, pairwise nucleotide divergences were calculated between the 21 ITS sequences obtained herein and the 10 sequences retrieved from GenBank that belong to the same sSA *Ilex* taxa. The values were calculated using the Kimura two-parameter model (Kimura, 1980) with DNAdist 3.5c program implemented in BioEdit (Hall, 1999). Additionally, ITS sequences were screened using BioEdit (Hall, 1999) for length variation, G + C content, and the presence of a structural motif in ITS 1 (i.e., GGCRY(4 to 7 n)GYGYCAAGGAA; Liu and Shardi, 1994). This conserved motif is often associated with part of the hairpin and loop structures that may serve as a critical signal in the enzymatic processing of the ribosomal RNA (Liu and Shardi, 1994). Furthermore, optimal and suboptimal folding structures, and associated free energy values (ΔG) in Kcal/mol, of ITS 1 and ITS 2 were recorded. Fold predictions were made at the Quickfold web server (Zuker, 2003) of the MFOLD program, version 3.1. In all cases, foldings were conducted at 37°C by use of a search within 5% of the thermodynamic optimality set. Because each sequence could adopt more than one folding structure, maximum and minimum ΔG values were recorded for each sequence, together with the number of all possible (optimal and suboptimal foldings) secondary structures. However, no attempts were made to study the RNA structures per se. These screenings were extended to the remaining representatives of *Ilex* and outgroup.

Phylogenetic and sensitivity analyses—Parsimony was chosen as the optimality criterion, as implemented in the program for Direct Optimization POY, version 3.0.11 (Gladstein and Wheeler, 1997; Wheeler and Gladstein, 2000; De Laet and Wheeler, 2003). To save computation time, as suggested in the POY user's manual and in Giribet (2001), the sequence data set was partitioned. The first partition comprises 257 to 265 bp, and the second partition has ca. 340 bp. Approximately 45 positions of the 5.8S gene, representing missing data in several sequences downloaded from the GenBank, were excluded from the analyses. A parameter space of two analytical variables was examined, i.e., gap cost and transversions/transitions (Tv/Ts) ratio cost. Precisely, we used gap costs = 1, 2, 4, 8 and 10, and Tv/Ts costs = 1 (equal weights), 2 (transversions receive twice the weight than transitions) and 4. All combinations of transformation costs used were in accordance with the triangle inequality (Wheeler, 1993, 1995). Each particular tree search was performed as follows: 50 independent heuristic searches starting from the best of five independent Wagner trees were conducted via subtree-pruning and regrafting (SPR) followed by TBR, retaining two trees per step, as specified in the following commands: “–norandomizeoutgroup –buildsperreplicate 5 –replicates 50 –maxtrees 2 –slop 2 –checkslop 5”. Commands “–slop” and “–checkslop” were employed to reduce error derived from the heuristic operations (Wheeler, 2002). Step matrices (i.e., transformation matrices) were specified in POY using the command “–molecularmatrix” with an argument for the given step matrix. Strict consensus trees were calculated using the command: “–poystriictconsensustreefile”. An implied alignment was generated in each case by the command: “–impliedalignment”. Approximate Bremer support values (Bremer, 1988) were calculated with a heuristic approximation procedure implemented in POY (command: “–bremer”). Finally, strict consensus trees and single trees were used to obtain 50% majority rule consensus tree. Implied alignments and trees were deposited at TreeBase (S996). We specifically used nodal stability as defined by Giribet (2003), i.e., the degree to which nodes are affected by variation in the analytical conditions. Robustness was used here-

TABLE 1. Number and relative abundance of AFLP monomorphic (m) and diagnostic (d) bands obtained per primer combination for selected *Ilex* taxa.

Primer combination	Total bands	m	m/total × 100 (%)	d	d/m × 100 (%)
E + AGG/M + CTT	224	28	12.5	12	43
E + ACT/M + CCA	210	57	27	30	53
E + AGC/M + CAT	193	42	22	21	50
E + ACC/M + CAG	209	58	28	31	53
Total	836	185	22	94	51

in as a measure of stability of nodes to fluctuation in parameters values across an analytical landscape (Ogden and Whiting, 2003).

Two sets of analyses were performed. First, we undertook a phylogenetic analysis of all ITS sequences, including those tentatively determined as paralogues and pseudogenes. To explore the influence of alignment parameters on the relative placement of these sequences, five different combinations of the two variables were employed. In particular, these combinations were: 111 (i.e., gap cost = 1, Tv = 1 and Ts = 1), 241, 421, 811 and 1021. Then, we performed a broader sensitivity analysis using only presumably active ITS sequences under 15 analytical conditions (namely, 111, 121, 141, 211, 221, 241, 411, 421, 441, 811, 821, 841, 1011, 1021 and 1041). Each of the 15 tree searches was performed as indicated earlier.

Combined analysis—The combined analysis involved AFLP data and ITS sequences. The sensitivity analysis was performed in POY as indicated previously, with the exception that a third file consisting of AFLP data per species was included in all calculations. The new data matrix for species was extracted from the original AFLP matrix for individuals and consisted of eight taxa and 578 characters. A band was considered present in a taxon (1) if it was recorded in more than 80% of the individuals, whereas it was considered absent in a taxon (0) if it was recorded in less than 20% of the individuals. Otherwise, bands were considered as polymorphic (0,1). The weights of AFLP changes were set so that nucleotide characters were never upweighted relative to AFLP data. The computer program POY allows nonsequence character changes to be simultaneously optimized with molecular data on multiple trees.

TABLE 2. Number of AFLP monomorphic (m) and diagnostic (d) bands recognized per *Ilex* species.

Species	m	d	d/m × 100 (%)
<i>Ilex argentina</i>	58	19	33
<i>I. brasiliensis</i>	60	13	22
<i>I. brevicuspis</i>	40	19	47.5
<i>I. dumosa</i>	13	8	61.5
<i>I. integerrima</i>	74	13	17.5
<i>I. paraguayensis</i>	38	17	44.7
<i>I. theezans</i>	69	5	7

RESULTS

AFLP analysis—The four primer combinations generated 836 scoreable bands, of which 22% were monomorphic (Table 1). All 120 individuals tested shared a single constant band. Primer combinations E + ACT/M + CCA and E + ACC/M + CAG yielded a nearly identical number of monomorphic and diagnostic bands. Among the seven species included in this study, 94 diagnostic bands were distinguished, which accounted for more than 50% of the monomorphic fragments recorded. The proportion of monomorphic and diagnostic bands detected varied among the species (Table 2). *Ilex integerrima* had the highest number of monomorphic bands, whereas the highest number of diagnostic bands was obtained for *I. argentina* and *I. brevicuspis*. In the box plot in Fig. 1, *I. argentina* gave the highest mean number of AFLP bands (mean = 175, SD = 29.9), whereas the lowest number of bands was found in *I. brevicuspis* (mean = 124, SD = 6.1).

Intra- and interspecific distance values derived from AFLP data are presented in Table 3. Excluding *I. dumosa*, for which the highest intraspecific distance value was obtained (0.1223, SD = 0.0408), remaining taxa had similar values. When interspecific mean distances were compared, *I. integerrima* and *I. theezans* appeared to be the least distant pair of taxa (0.1014, Table 3).

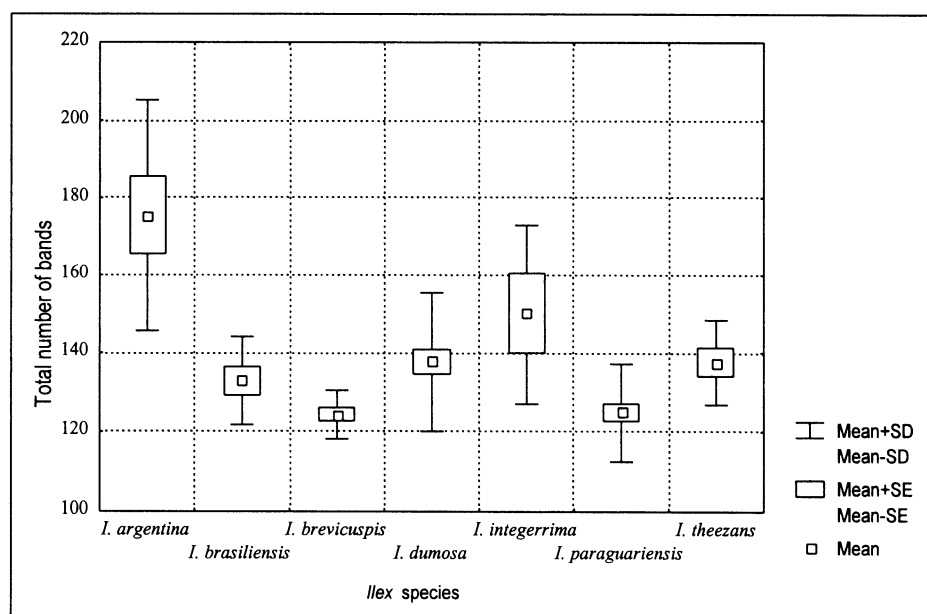
Fig. 1. Comparison of mean number of AFLP bands obtained for selected *Ilex* species.

TABLE 3. Intra- (diagonal line in bold) and interspecific mean distance values (Nei and Li, 1979) for sSA *Ilex* taxa based on AFLP data.

Taxa	<i>I. argentina</i>	<i>I. brasiliensis</i>	<i>I. brevicuspis</i>	<i>I. dumosa</i>	<i>I. integerrima</i>	<i>I. paraguariensis</i>	<i>I. theezans</i>
<i>I. argentina</i>	0.0738						
<i>I. brasiliensis</i>	0.1582	0.0718					
<i>I. brevicuspis</i>	0.2061	0.1996	0.0633				
<i>I. dumosa</i>	0.241	0.2657	0.2570	0.1223			
<i>I. integerrima</i>	0.1801	0.1267	0.2079	0.2577	0.0739		
<i>I. paraguariensis</i>	0.2211	0.2394	0.2449	0.2303	0.2271	0.0692	
<i>I. theezans</i>	0.1616	0.1096	0.1869	0.2491	0.1014	0.2367	0.0624

Of the 836 binary characters, 749 were parsimony informative. The heuristic search resulted in 148 most parsimonious trees (MPT) with a length of 5260 (CI = 0.1561, RI = 0.66). The strict consensus tree derived from AFLP data showed that all individuals of the same taxon grouped accordingly (Fig. 2). Accessions of *I. theezans* yielded a coherent group (BV, bootstrap value; BV = 93%), closely related to the clade *I. integerrima* (BV = 95%). Within *I. brasiliensis* (BV = 66%), two subclades were distinguished, one comprised accessions from Argentina and Paraguay (BV = 83%), and the other subclade involved individuals from Brazil (BV = 82%). In contrast, specimens of *I. argentina* formed a highly supported clade (BV = 97%). The clade *I. brevicuspis* (BV = 100%) showed several internal nodes unresolved and poorly supported. Within the *I. dumosa* clade (BV = 99%), some geographical structuring was detected. The first subclade (BV < 50%) involved individuals from Argentina and Paraguay representing both varieties of *I. dumosa* (*I. dumosa* var. *dumosa* and *I. dumosa* var. *guaranina*, the latter indicated with an arrow in Fig. 2), while the other subclade comprised solely Brazilian accessions (BV = 68%). Regarding *I. paraguariensis* (BV = 100%), the majority of Argentine, Brazilian, and Paraguayan accessions of mate appeared intermingled (BV = 96%); still, most internal relationships were poorly supported. Commercial lines of mate (SI-16, SI-19, SI-167 and G-18) formed a clade with a specimen from Paraguay (BV = 59%). Another subclade involved lines SI-49, Y-383, and two accessions from Paraguay (BV = 66%). The relationship between *I. brasiliensis*, *I. integerrima*, and *I. theezans* was highly supported by the data (BV = 100%). Likewise, the relationship of *I. argentina* and *I. brevicuspis* to the aforementioned taxa was also well-supported by the data (BV = 91%).

ITS sequence studies—Nucleotide divergence calculations showed that in several cases sSA ITS sequences retrieved from GenBank were highly divergent from those obtained in this report (Table 4). For instance, similar divergence values (ca. 14–15%) were found for ITS sequences of *I. brasiliensis*, *I. integerrima*, and *I. pseudobuxus*. While *I. brevicuspis* AJ492663 diverged with our ITS data in ca. 11%, accession AJ492656 of *I. dumosa*, was ca. 21% divergent from *I. dumosa* AJ492657 and with the five sequences of *I. dumosa* obtained herein. On the other hand, for *I. theezans*, *I. microdonta*, *I. brevicuspis* AJ492662, and *I. dumosa* AJ492657, the range of nucleotide variation detected was approximately under 5%. In all cases, divergence values calculated among co-specific sequences obtained in this study were lower than 2%. The length, G + C content, free energy values of the RNA transcripts, and the presence of a conserved structural motif were also evaluated in ITS 1 and ITS 2 of sSA and worldwide distributed *Ilex* taxa (Table 5). For *I. microdonta* and *I. theezans*,

no distinct differences were observed between our sequences and those from GenBank. Instead, major differences were found between our nucleotide data and sequences *I. brasiliensis* AJ492661, *I. brevicuspis* AJ492663, *I. dumosa* AJ492656, *I. integerrima* AJ492664, and *I. pseudobuxus* AJ492660. These downloaded data showed a markedly low G + C content and a considerable reduction in stability of the most probable secondary structure, both for ITS 1 and ITS 2. *Ilex brevicuspis* AJ492662 and *I. dumosa* AJ492657 were not completely comparable to other co-specific data due to a difference in length. Unfortunately, these sequences were incomplete lacking at least, 60 bp from the 5' end of ITS 1. Notwithstanding, their ITS 2 had comparable values to our sequences in all factors considered. Excluding these incomplete sequences, no significant differences in length were encountered. Among North American and East Asian sequences, some variations in the screened parameters were found (Table 5). In particular, the values obtained for ITS 2 of *I. kiusiana* make us question its functionality. The lack of additional sequences impeded us from making a final decision.

The universal ITS 1 core motif of Liu and Schardl (1994) was screened in all 55 sequences, and some variants were detected. Particularly in *Ilex*, 13 variants of this sequence were recognized and arbitrarily named (Table 6). The types B and D were the most frequent ones (21/53 and 11/53, respectively). This scrutiny revealed that all six *I. dumosa* sequences (AY183479, AY183480, AY183481, AY183488, AY183489, and AJ492657) had a different motif that lacks the final adenine. As to the GenBank data (*I. brasiliensis* AJ492661, *I. brevicuspis* AJ492663, *I. dumosa* AJ492656, *I. integerrima* AJ492664, and *I. pseudobuxus* AJ492660), these exhibited a motif that substantially departs from the core sequence, contributing to a less stable secondary structure (Tables 5 and 6). Two of these (*I. dumosa* AJ492656 and *I. brevicuspis* AJ492663) were previously reported as paralogous (Manen et al., 2002).

Phylogenetic and sensitivity analyses—The input data set for the phylogenetic analysis comprised 53 ingroup sequences and two outgroup, also including putative divergent paralogues and pseudogenes. Analytical conditions 241, 421, and 1021 yielded a single MPT of length 2437, 2796, and 4707, respectively (not shown), whereas conditions 811 and 111 rendered two and three MPTs of length 3617 and 1346, respectively (not shown). The 50% majority rule consensus tree of Fig. 3, condenses the information display in all MPT encountered. Regarding sSA taxa, sequences of *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, *I. microdonta*, *I. pseudobuxus*, *I. taubertiana*, *I. theezans*, *I. microdonta* AJ492665, and *I. theezans* AJ492666 appeared related in all the conditions assayed. However, under conditions 111 and 241, se-

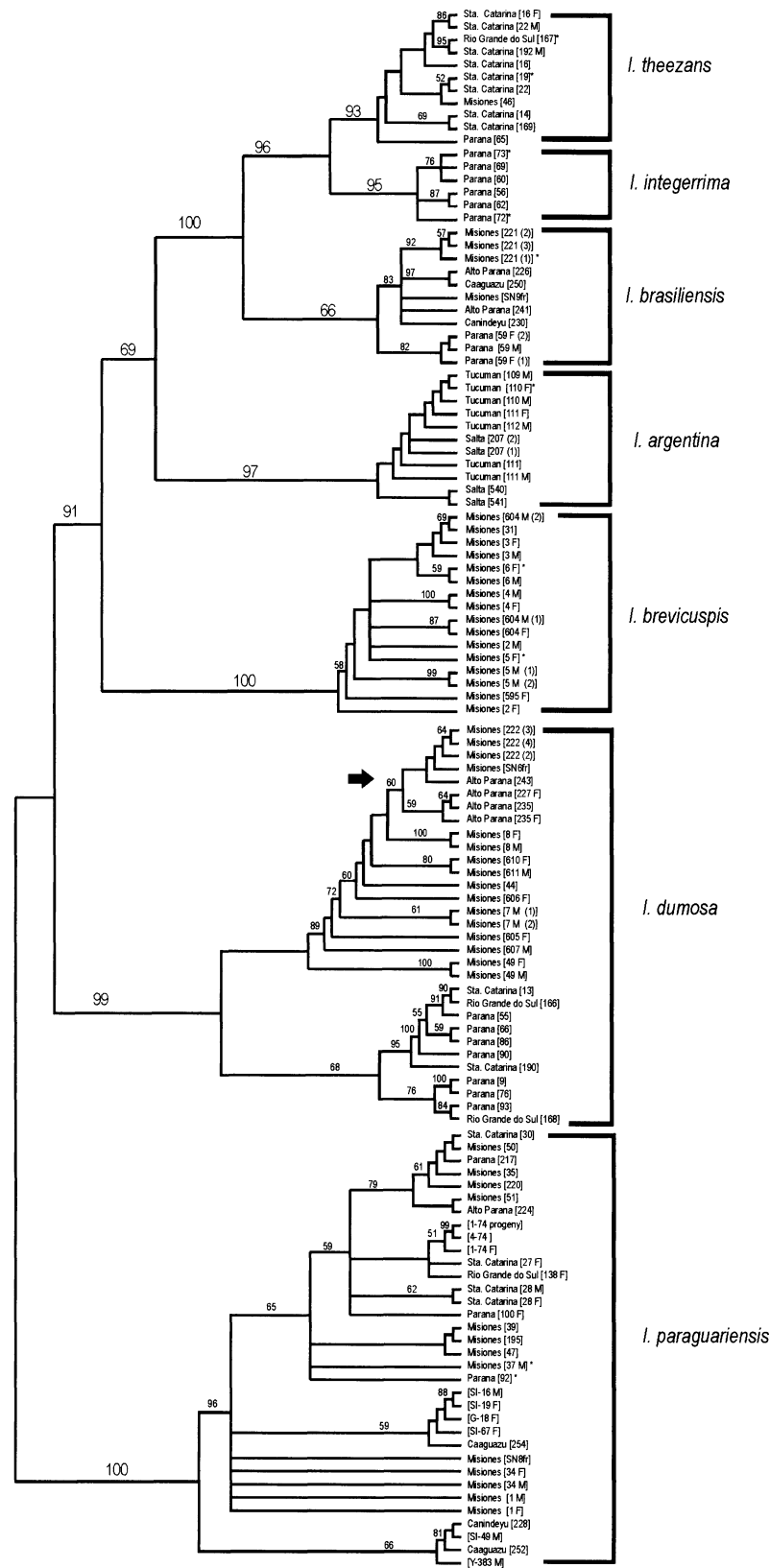


Fig. 2. Maximum parsimony strict consensus tree of selected *Ilex* species derived from AFLP data. Numbers above branches indicate bootstrap values (1000 replicates). Accessions, as presented in Appendix 1 (see Supplemental Data accompanying the online version of this article), are indicated between brackets for clearness. F, stands for female; M, stands for male. The arrow indicates terminals belonging to *I. dumosa* var. *guaranina*. *, indicates specimens used for sequencing.

TABLE 4. Intraspecific nucleotide divergence between *Ilex* species derived from Kimura two-parameter model (Kimura, 1980). Each individual sequences is identified by its corresponding GenBank accession number. Bolded sequences were obtained in this study.

<i>Ilex argentina</i>	AY183482	AJ492655			<i>I. brasiliensis</i>	AY183490
AY183483	0.0015	0.0946			AJ492661	0.1426
AY183482		0.0927				
<i>I. integerrima</i>	AY183477	AJ492664			<i>I. microdonta</i>	AY183484
AY183478	0.0078	0.1409			AY183485	AJ492665
AY183477		0.154			AY183484	0.0015
						0.0271
						0.0286
<i>I. paraguariensis</i>	AY183491				<i>I. pseudobuxus</i>	AY183486
AY183492	0.0092				AJ492660	0.1494
<i>I. brevicuspis</i>	AY183494	AY183496	AJ492662	AJ492663		
AY183495	0.0185	0.0184	0.0262	0.1249		
AY183494		0.0031	0.0068	0.1146		
AY183496			0.005	0.109		
AJ492662				0.1072		
<i>I. theezans</i>	AY183497	AY183493	AJ492666			
AY183498	0.0127	0.0338	0.0515			
AY183497		0.0159	0.0222			
AY183493			0.0197			
<i>I. dumosa</i>	AY183488	AY183481	AY183480	AY183479	AJ492657	AJ492656
AY183489	0.0046	0.0082	0.0113	0.0123	0.0117	0.215
AY183488		0.0049	0.0081	0.0092	0.0067	0.2114
AY183481			0.0086	0.0098	0.0108	0.2276
AY183480				0.008	0.0088	0.2059
AY183479					0.005	0.2143
AJ492657						0.2073

quence *I. brevicuspis* AJ492662 appeared to be related to our allele with a low Bremer support value (2 in condition 111 and 4 in condition 241, not shown). Neither the two sequences of *I. theezans* nor both sequences of *I. microdonta* formed monophyletic groups. However, both *I. theezans* appeared as sister in 40% of the parameters sets, while both *I. microdonta* appeared as sister in 100% of the conditions. ITS sequences of *I. brasiliensis* AJ492661, *I. brevicuspis* AJ492663, *I. integerrima* AJ492664, and *I. pseudobuxus* AJ492660 formed a clade under the five analytical conditions explored. This clade received Bremer support values that ranged from 36 to 81, depending on the analytical conditions. These values represented the second to third highest values compared to all values obtained with each condition. Hence, these sequences never formed monophyletic groups with their corresponding counterparts. Although the placement of *I. dumosa* AJ492656 varied with the parameters, it never appeared monophyletic with the remaining *I. dumosa*. On the other hand, *I. dumosa* AJ492657 appeared as sister to both *I. dumosa* obtained herein, in 40% of the conditions, whereas it was placed as sister to *I. brevicuspis* AJ492662 in 60% of the conditions. In these latter cases, the Bremer support values were extremely high, ranging from 94 to 389. *Ilex argentina* AJ492655 never appeared related to its counterpart, while the placement of *I. paraguariensis* was highly dependent on parameter sets.

As a result, several features suggest that accessions *I. brasiliensis* AJ492661, *I. brevicuspis* AJ492663, *I. dumosa* AJ492656, *I. integerrima* AJ492664, and *I. pseudobuxus* AJ492660, could be considered as pseudogenes: the high nucleotide divergence with co-specific data, the low G + C content, the departure from the conserved motif in ITS 1, the less thermodynamic stability in the RNA structure, and their basal position on phylogenetic trees with respect to sSA counterparts. Therefore, only *I. theezans* AJ492666, *I. microdonta* AJ492665, and *I. argentina* AJ492655, retrieved from GenBank, were further used as alternative functional ITS types.

The phylogenetic analysis restricted to active ITS sequences, comprised 46 ingroup sequences and two outgroup. Strict consensus trees and single MPTs obtained under each combination of gap cost and Tv/Ts ratio cost are shown in Figs. 4–18. The 50% majority rule consensus tree (Fig. 19, on the left), condenses the information display in all MPT encountered. The overall topology of the trees was not altered by the exclusion of the pseudogenes, although its resolution increased. Regarding sSA taxa, the clade composed by *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, *I. microdonta*, *I. pseudobuxus*, *I. taubertiana*, and *I. theezans*, including sequences *I. microdonta* AJ492665 and *I. theezans* AJ492666 (hereafter, sSA clade) appeared in 100% of the conditions assayed. This clade received Bremer support values that ranged from 16 to 58, depending on the analytical condition. These values represent the third to fifth highest values compared to all values obtained in each condition. As before, the pairs (*I. brasiliensis* + *I. theezans*) and (*I. integerrima* + *I. theezans* AJ492666) were found throughout the explored weights, although both subclades received Bremer supports from 7 to 24, which represent the seventh to less than tenth highest values. Similarly, the subclade (*I. microdonta* + *I. pseudobuxus* + *I. microdonta* AJ492665) appeared in 100% of the analyses, but its Bremer support was either the tenth highest value or lower. The formation and placement of subclade (*I. brevicuspis* + *I. taubertiana*) was dependent on analytical conditions; it was found in 80% of the conditions, and it formed a monophyletic group with (*I. microdonta* + *I. pseudobuxus* + *I. microdonta* AJ492665) in only 53% of the conditions. The sister sequence to the sSA clade, in 67% of the parameters sets assayed, was represented solely by Asian *I. pedunculosa*, but with low Bremer support values. Solely under analytical condition 811 (Fig. 13), *I. paraguariensis*, both varieties of *I. dumosa*, and East Asian *I. crenata* and *I. mutchagara*, appeared as sister group to the sSA clade, with a Bremer support of 29 (seventh highest

TABLE 5. Results from ITS 1 and ITS 2 sequence comparisons.

Taxa*	ITS 1					Consensus sequence**	ITS 2				
	Length	G + C (%)	ΔG optimal (Kcal/mol)	ΔG suboptimal (Kcal/mol)	Structures		Length	G + C (%)	ΔG optimal (Kcal/mol)	ΔG suboptimal (Kcal/mol)	Structures
SOUTH AMERICA											
<i>Ilex argentina</i>	260	59.23	-99.7	-94.8	8	B	240	56.67	-90.4	-87.6	4
<i>I. argentina</i> AJ492655	262	59.54	-104.7	-100	9	D	248	59.68	-87.4	-83.3	4
<i>I. brasiliensis</i>	262	62.6	-103.8	-98.9	12	B	242	58.68	-104.3	—	1
<i>I. brasiliensis</i> AJ492661	259	50.97	-84.6	-80.7	6	J	245	53.06	-87.9	-83.6	10
<i>I. brevicuspis</i>	263	61.22	-104	-98.8	14	B	242	57.85	-95	-90.3	4
<i>I. brevicuspis</i> AJ492662	203	63.55	-72.9	—	1	B	248	58.06	-97.5	-93.7	4
<i>I. brevicuspis</i> AJ492663	259	49.03	-74.3	-70.6	18	L	242	54.55	-84.6	-80.7	10
<i>I. dumosa</i> var. <i>dumosa</i>	263	60.84	-112.1	-106.5	8	G	241	58.09	-94.1	-90.5	11
<i>I. dumosa</i> var. <i>guaranina</i>	263	61.6	-114.7	-109.6	7	G	241	58.51	-99.9	-95	6
<i>I. dumosa</i> AJ492657	205	63.41	-76.6	-73.1	4	G	247	58.7	-102.1	-97.1	5
<i>I. dumosa</i> AJ492656	251	45.42	-60.9	-58.2	7	K	238	52.52	-70.6	-67.4	11
<i>I. integerrima</i>	264	62.12	-105.5	-100.4	10	B	240	58.33	-88.4	-84.9	3
<i>I. integerrima</i> AJ492664	259	51.74	-86.4	-82.4	13	J	244	52.46	-87.6	-83.5	11
<i>I. microdonta</i>	265	62.64	-106.2	-101	18	B	244	59.02	-100.7	-96.4	4
<i>I. microdonta</i> AJ492665	265	61.89	-108.5	-103.4	10	B	248	60.89	-102.5	-97.7	6
<i>I. paraguayensis</i>	269	61.71	-114.9	-109.2	18	F	240	59.17	-101.1	-96.7	2
<i>I. pseudobuxus</i>	265	63.77	-110	-104.7	11	B	242	60.74	-96.5	-91.7	9
<i>I. pseudobuxus</i> AJ492660	258	48.45	-82.4	-78.4	9	M	245	51.84	-83.8	-80.2	9
<i>I. taubertiana</i>	263	58.56	-97	-92.6	11	A	244	57.38	-98	-95.6	3
<i>I. theezans</i>	263	60.08	-96.5	-92.5	3	B	242	57.85	-94.4	-90.5	4
<i>I. theezans</i> AJ492666	264	62.88	-110	-104.7	20	B	249	59.04	-107.9	-102.7	2
EAST ASIA											
<i>I. mitis</i>	260	61.54	-106.8	-102.1	7	B	244	56.56	-97.6	-97.4	2
<i>I. beecheyii</i>	262	59.92	-113.3	-108.7	7	B	223	59.19	-89.3	-84.9	8
<i>I. buergeri</i>	261	58.62	-104.5	-99.5	7	C	238	59.24	-84.3	-80.1	16
<i>I. cornuta</i>	261	61.3	-115.7	-110	12	B	237	58.23	-78.7	-75.3	14
<i>I. crenata</i>	260	60.77	-103.6	-98.5	18	A	239	62.34	-96.7	-92.2	23
<i>I. dimorphophylla</i>	261	62.07	-118.8	-113.3	10	B	237	59.07	-87.9	-83.8	15
<i>I. integra</i>	263	56.27	-100.4	-95.4	12	H	224	58.04	-91	-86.9	12
<i>I. kiusiana</i>	262	59.16	-105.3	-100.7	7	C	228	53.07	-70.2	-66.9	5
<i>I. latifolia</i>	260	59.62	-104.5	-99.3	8	D	238	57.98	-87.9	-85.5	2
<i>I. liukiensis</i>	256	59.38	-97.5	-92.8	16	E	237	57.38	-78.9	-75	13
<i>I. macropoda</i>	257	59.14	-101.2	-96.9	8	C	239	55.65	-82.2	-78.7	13
<i>I. maximowicziana</i>	260	60.77	-104.2	-100.3	13	B	237	59.49	-88.1	-83.8	13
<i>I. mertensii</i>	263	60.08	-113.8	-108.4	9	A	223	59.19	-89.3	-84.9	8
<i>I. micrococca</i>	259	59.07	-97.9	-93.7	14	C	243	59.67	-100.8	-96	4
<i>I. mutchagara</i>	262	61.45	-107.6	-102.4	11	A	239	64.44	-101.1	-96.1	15
<i>I. pedunculosa</i>	263	60.46	-113.8	-108.4	8	B	242	57.44	-101.6	-97	4
<i>I. percoriacea</i>	263	60.08	-113.8	-108.4	9	A	223	59.19	-89.3	-84.9	8
<i>I. rotunda</i>	260	63.85	-110	-104.8	11	B	248	64.92	-113.6	-108.6	7
<i>I. rugosa</i>	256	57.81	-94.4	-90.4	15	B	238	55.04	-87.8	-83.9	7
<i>I. serrata</i>	258	59.69	-96.8	-92.3	14	D	240	55	-92.4	-88	10
<i>I. warbugii</i>	260	59.62	-106.1	-101.5	5	D	240	58.33	-83.6	-79.5	19
<i>I. yunannensis</i>	258	59.69	-102.8	-97.7	14	D	241	57.68	-94.3	-90.8	4
<i>I. anomala</i>	258	54.26	-92.7	-88.3	9	I	231	55.41	-83.7	-81.9	5
<i>I. canariensis</i>	261	65.13	-120.8	-115.2	11	B	247	63.16	-105.9	-101.6	10
NORTH AMERICA											
<i>I. amelanchier</i>	257	59.53	-100.3	-96.8	9	B	242	52.89	-83.9	-79.9	6
<i>I. cassine</i>	262	59.54	-104.7	-100	9	D	249	60.24	-91.7	-87.2	8
<i>I. collina</i>	260	58.08	-107.6	-102.9	9	D	234	54.7	-89.1	-84.9	8
<i>I. decidua</i>	260	58.08	-101.4	-97.3	10	D	234	56.84	-96.2	-92.1	5
<i>I. glabra</i>	261	58.62	-94.1	-89.5	17	D	243	58.02	-101.8	-96.9	14
<i>I. mucronata</i>	256	59.38	-102.6	-97.8	9	B	242	54.13	-87.6	-83.8	5
<i>I. opaca</i>	259	58.69	-105.7	-100.5	14	D	243	59.26	-91.5	-88.1	3
<i>I. vomitoria</i>	260	58.46	-102.7	-98.6	9	D	234	56.84	-96.2	-92.1	5
<i>Helwingia japonica</i>	264	56.44	-93.8	-89.2	10	X	227	53.74	-86.3	-82.8	7
<i>H. chinensis</i>	265	56.23	-94.2	-89.6	15	X	227	53.3	-85.6	-82.1	6

* Sequences in bold were obtained in this study (see Materials and Methods, and Supplemental Data accompanying the online version of this article). Sequences identified with GenBank accession numbers are those of Manen et al. (2002). For remaining sequences, see Supplemental Data accompanying the online version of this article. **, see Table 6.

TABLE 6. ITS 1 motives of Liu and Schradl (1994) found in 55 ITS 1 *Ilex* sequences and outgroup sequences (for detailed distribution of each type, see Table 5).

	Type
GGCGCTGTTTGGCGCAAGGAA	A
GGCGCTGTCTGCGCAAGGAA	B
GGCGCTATTTGGCGCAAGGAA	C
GGCGCTATCTGCGCAAGGAA	D
GGCGCTATCTGCG GAACCAT	E
GGCGCTGTCTGTGCAAGGAA	F
GGCGCTGTATGCGCAAGGACCCAT	G
GGCGCTTTTTGGCGCAAGGAA	H
GGCGCTATATGTGCAAGGAA	I
GG T GTATCTGTGCAAGGAA	J
AA CACTATCT ACA CCAAGGAA	K
GGCGTATCTGT A CCAAGGAA	L
CGACG TATCTGTGCA A AGGAA	M
GGCGCA AA AGCGCAAGGAA	X

Note: Bold and italics = changed bases compared to consensus motif of Liu and Schardl (1994) see Materials and Methods.

value). In the analytical space explored, *I. dumosa* varieties consistently yielded a clade with the second highest Bremer value (18–99). The varieties appeared more closely related to East Asian *I. crenata* and *I. mutchagara* (73% of the conditions) than to other sSA taxa, excluding *I. paraguayensis*. The latter taxa appeared basal in relation to remaining ingroup taxa, in eight of the 15 parameters sets assayed, whereas under six analytical conditions, it was closely allied to *I. dumosa*, *I. crenata* and *I. mutchagara*.

Concerning North American taxa, the ITS sequences of *I. collina*, *I. decidua*, and *I. vomitoria* were related across all the analytical space, although with variable Bremer support (from 3 to 54), i.e., values ranging from the fourth to less than 10th highest. The pair of taxa *I. cassine* and *I. opaca* formed a clade with *I. argentina* AJ492655 in 100% of the conditions, with Bremer support values within the third and seventh highest (13–49). *Ilex amelanchier* and *I. mucronata* appeared monophyletic in 80% of the conditions. The location of *I. glabra* varied considerably under different analytical conditions. Likewise, the placement of Hawaiian *I. anomala* relies on analytical conditions, appearing basal to several East Asian taxa in 53% of the conditions.

The Asian taxa formed two stable groups in all of the analyses performed. One was composed of sequences of *I. beecheyii*, *I. integra*, *I. kiusiana*, *I. mertensii*, and *I. percoriacea*, and had variable Bremer support values ranging from 5 to 87 (second to ninth higher values). Instead, the other stable clade formed by *I. cornuta*, *I. dimorphophylla*, and *I. maximowicziana* had lower Bremer support values. Whereas, Asian taxa *I. buergeri*, *I. latifolia*, *I. liukuensis*, and *I. warburgii* appeared monophyletic in 93% of the conditions. The position of the remaining Asian taxa (i.e., *I. macropoda*, *I. micrococca*, *I. rotunda*, *I. serrata*, and *I. yunnanensis*) varied with alignment parameters. ITS sequences of African *I. mitis* and Macaronesian *I. canariensis* formed a clade in 60% of the parameter space.

Combined analysis—Additional analyses were carried out combining AFLP data for eight taxa and nucleotide sequences of active ITS alleles. A 50% majority rule consensus tree is shown as a summary of the sensitivity analysis (Fig. 19, on the right). The lengths of strict consensus trees and single MPT

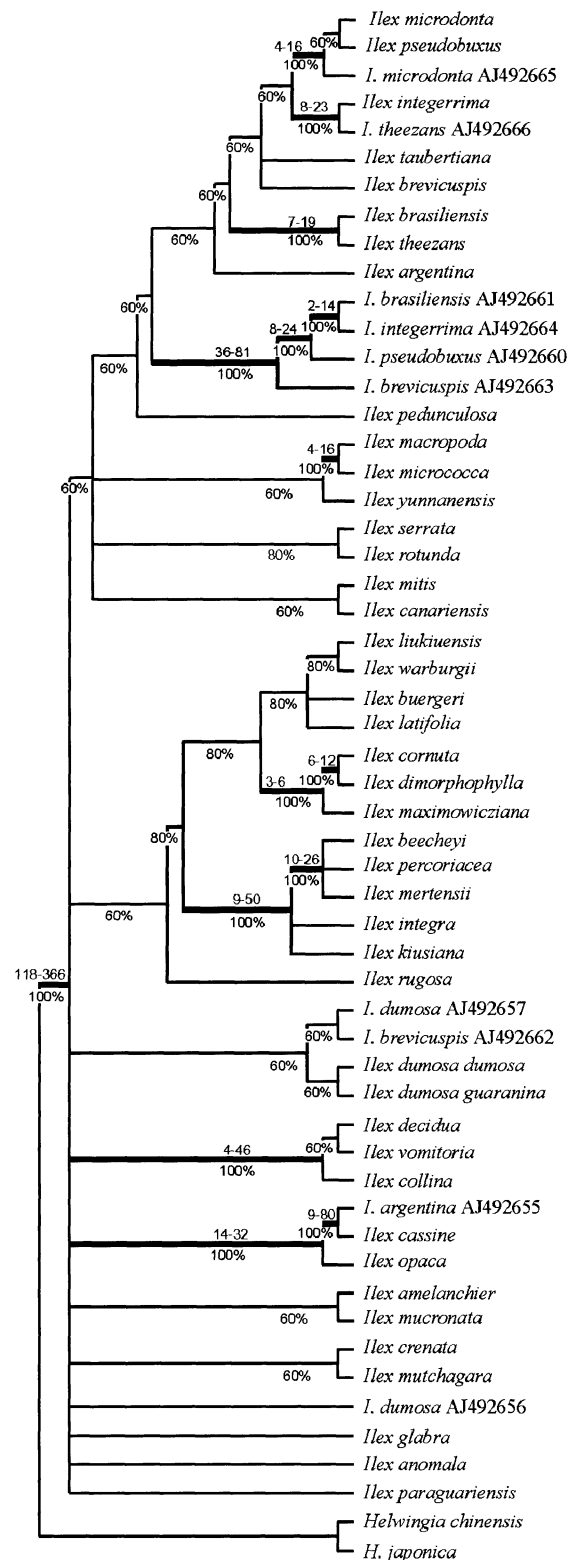
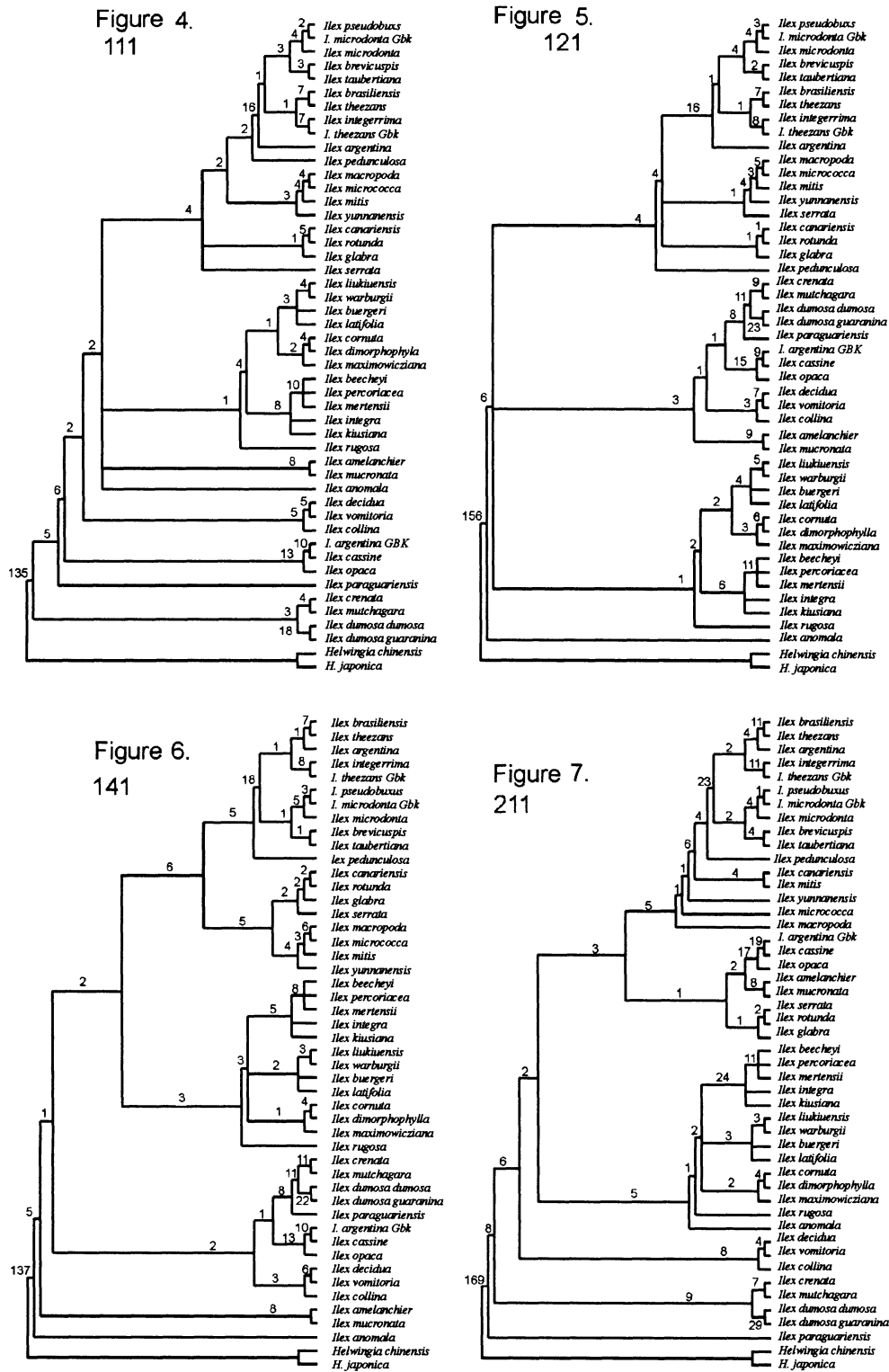
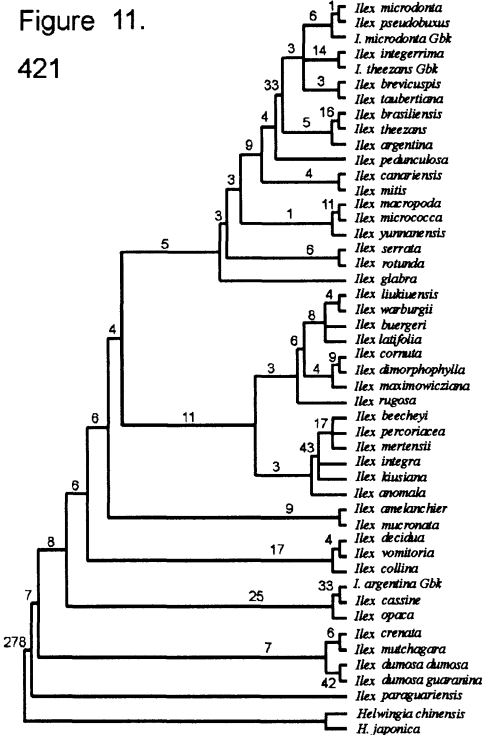
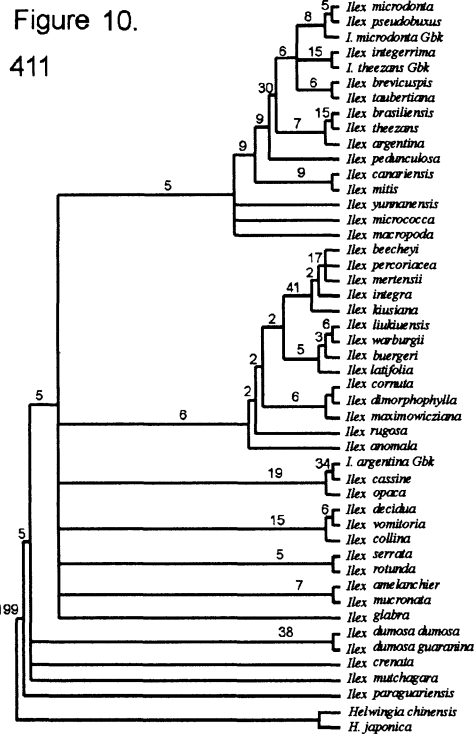
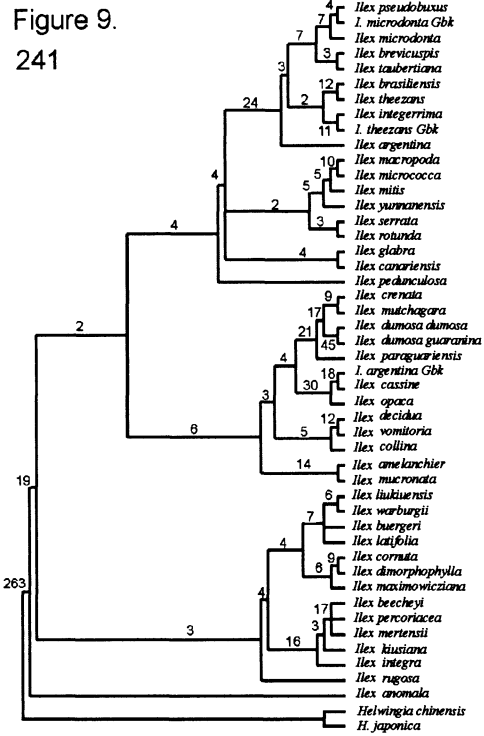
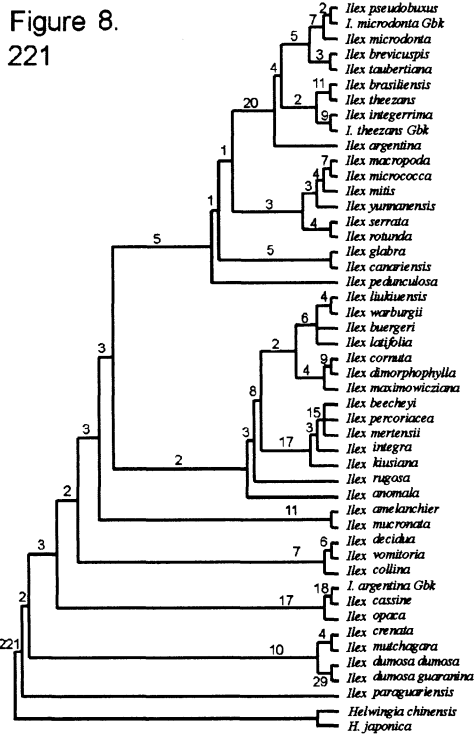


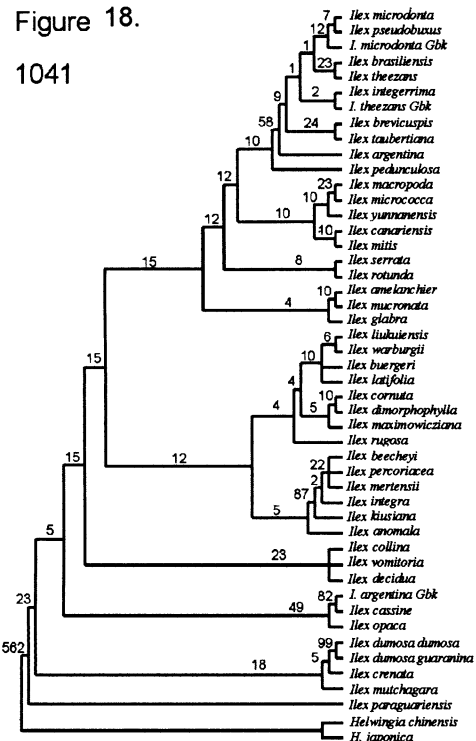
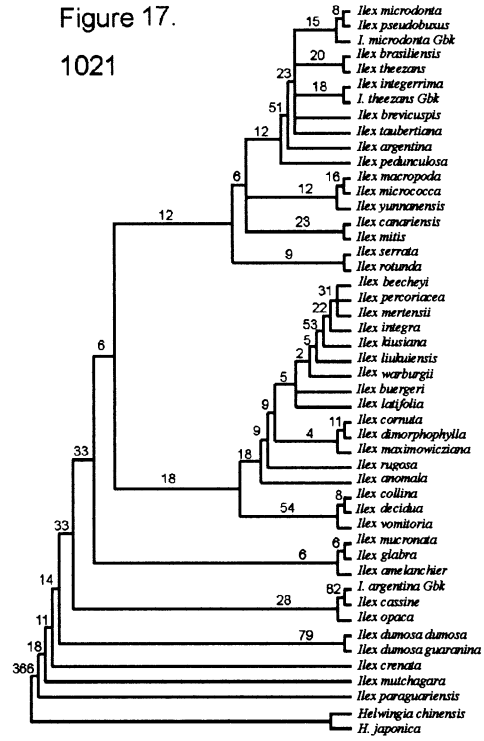
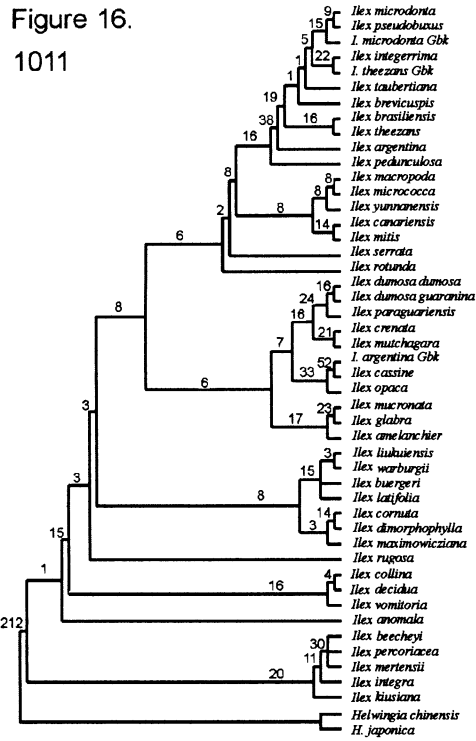
Fig. 3. Fifty percent majority rule consensus tree derived from the sensitivity analysis of *Ilex* ITS sequences, including divergent copies. Numbers below branches indicate percentage of formation of the node. Numbers above branches represent ranges of Bremer support values. Clades found throughout the parameter space explored are marked with thick lines.



Figs. 4–7. Results from the sensitivity analyses of *Ilex* taxa via direct optimization of ITS sequences. 4. Strict consensus tree of four most parsimonious trees (MPT) of length 1013, under parameters 111. 5. Strict consensus tree of eight MPT of length 1174, under parameters 121. 6. Strict consensus tree of 11 MPT of length 1297, under parameters 141. 7. Strict consensus tree of three MPT of length 1301, under parameters 211. *Figure abbreviations:* The transformation matrix stands for the step matrix used in each search (i.e., 241: gap cost = 2, Tv cost = 4, Ts cost = 1). Gbk, indicates sSA sequences retrieved from GenBank and discussed in the text.



Figs. 8–11. Results from the sensitivity analyses of *Ilex* taxa via direct optimization of ITS sequences. **8.** Single MPT of length 1533, obtained under parameter 221. **9.** Strict consensus tree of two MPT of length 1842, under parameters 241. **10.** Strict consensus tree of four MPT of length 1797, under parameters 411. **11.** Strict consensus tree of six MPT of length 2083, under parameters 421.



Figs. 16–18. Results from the sensitivity analyses of *Ilex* taxa via direct optimization of ITS sequences. 16. Single MPT of length 3023, obtained under parameter 1011. 17. Strict consensus tree of two MPT of length 3491, under parameters 1021. 18. Strict consensus tree of two MPT of length 4127, under parameters 1041.

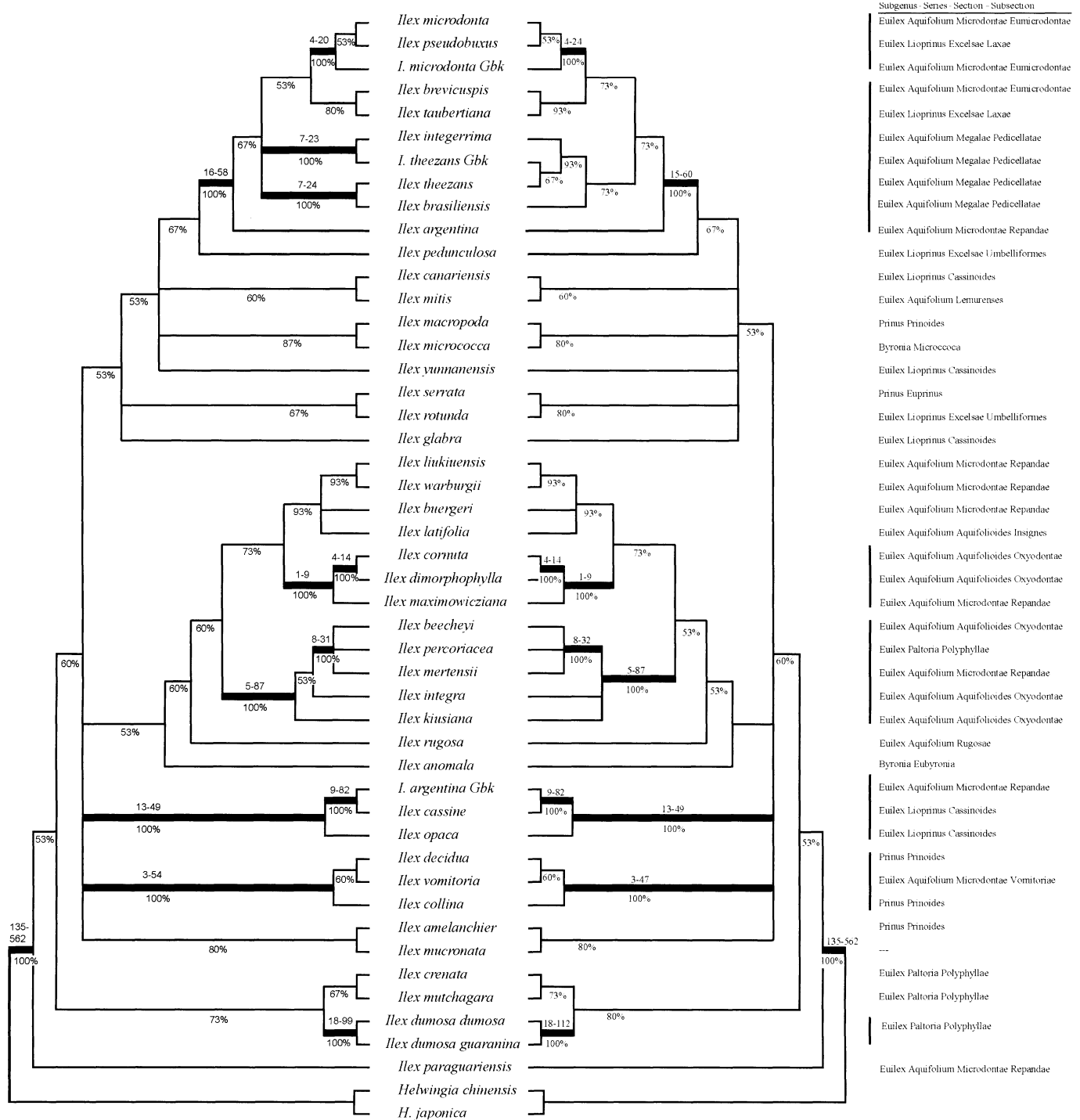


Fig. 19. Fifty percent majority rule consensus tree derived from the sensitivity analyses of ITS sequences alone (on the left) and from the combined analysis (on the right). Numbers below branches indicate percentage of formation of the node. Numbers above branches represent ranges of Bremer support values. Clades found throughout the parameter space explored are marked with thick lines. Gbk indicates sSA sequences downloaded from the GenBank and discussed in the text. Loesener's subgeneric disposition is presented.

obtained, ranged from 1356 to 7479. When compared with the ITS alone analysis, we observed that in 10/15 of the explored parameter space, the number of MPTs found remained unchanged; in 4/10 situations, the number of MPT was lower than in the ITS analysis alone; and only in one case, was the number of MPT higher (condition 411 rendered 10 MPT when combining and 4 in ITS alone). In general, the combined to-

polo was congruent with topology of ITS sequence alone, as were Bremer support values for stable nodes. Main differences were found within the sSA clade, though some Asian and North American taxa showed alternative placement. Within sSA, the relationship between *I. brasiliensis*, *I. integerrima*, and *I. theezans* was resolved in the combined analysis as with AFLP data alone. In contrast, the combination of data did not

resolve the relationship between both *I. microdonta* alleles (which appeared monophyletic in 47% of the conditions) and *I. pseudobuxus*. The placement of *I. argentina* appeared here slightly less dependent on parameters. Similarly, the addition of AFLP data slightly stabilized the relationship of *I. dumosa* with *I. crenata* and *I. mutchagara*. Notwithstanding, the data combination did not help in stabilizing the parameter-dependent position of *I. paraguariensis*.

Neither the analysis of ITS data alone nor the combined analysis, yielded trees that completely agreed with Loesener's taxonomic classification. Although the trees derived from ITS analysis under parameter sets 221 (Fig. 8), 441 (Fig. 12), and 821 (Fig. 14) had the highest number of clades in concordance with current classification, none of the supraspecific ranks appeared monophyletic.

DISCUSSION

AFLP analyses—The AFLP technique is highly efficient for detecting polymorphisms of restriction fragments by PCR amplification. The difficulty in identifying homologous markers, the potential non-independence of bands and the dominant nature of AFLP markers, are the main disadvantages of the technique (Mueller and Wolfenbarger, 1999). However, AFLP allows the simultaneous screening of many different high-resolution markers randomly distributed throughout the genome.

The combinations of selective AFLP primers employed in this study allowed the recognition of several diagnostic bands within each of the seven species assayed. These would be useful for molecular identification of the plants, particularly at early developmental stages. A relationship between the number of AFLP bands scored for each species and its ploidy level was detected, even though it was not statistically tested (Fig. 1). For instance, *I. argentina* had the most complex banding pattern, and it is the sole tetraploid species ($2n = 80$) detected in the region (Barral et al., 1995).

Our AFLP analysis revealed that individuals belonging to the same morpho-species formed coherent clades (with moderate to strong bootstrap support). Therefore, seven clades were distinguished corresponding to each of the species surveyed. The species *I. brasiliensis*, *I. integerrima*, and *I. theezans* had a strongly supported relationship. The branching order in which *I. argentina* and *I. brevicuspis* associate with *I. brasiliensis*, *I. integerrima*, and *I. theezans* was not undoubtedly defined by the data; nevertheless, the monophyly of the five species was very well-supported. Instead, the relationship to *I. dumosa* and *I. paraguariensis* of the aforementioned species was not clearly established.

An important aspect of our analysis is the observation that AFLP data do not confirm the traditional distinction, established by Loesener (1901), between *I. dumosa* var. *dumosa* and *I. dumosa* var. *guaranina*, which was based on the inflorescence type. It seems possible that morphological features used to distinguish and define the varieties (Loesener, 1901; Giberti, 1994a) are not associated with the differences observed herein. Instead, our results suggest that within *I. dumosa*, Argentine and Paraguayan populations were derived from a common ancestral population a posteriori of the divergence from Brazilian population. A similar association was observed within *I. brasiliensis*. On the contrary, no geographical pattern was noted within *I. paraguariensis*, for which internal nodes appeared, in general, poorly supported and resolved. For this species, calculated mean intraspecific distance

was among the lowest values obtained. Similarly, Gauer and Cavalli-Molina (2000), using RAPD markers, found low divergence among natural populations of *I. paraguariensis* from Brazil. It seems reasonable that the arguments used by those authors (the occurrence of gene flow favored by fruit-eating birds or recent fragmentation caused by deforestation and/or reduction of natural forests) are responsible for the results of both studies.

ITS sequence studies—The nucleotide divergence comparison revealed that in several cases, divergence between sequence data obtained in this study and sequence data downloaded from GenBank, far exceeded the values reported in the literature, usually below 5% (Baldwin et al., 1995). Results presented here, as a rule, do not support the view that the ITS sequences retrieved from GenBank (those of Manen et al., 2002), and our ITS sequences reflect intraspecific polymorphism. In general, the sSA ITS sequences of Manen et al. (2002) were highly divergent from those obtained herein, have lower G + C content, and were less thermodynamically stable. Previously, Manen et al. (2002) reported two of their South American ITS sequences as paralogous (*I. dumosa* AJ492656 and *I. brevicuspis* AJ492663). In this study, we gathered solid evidence indicating that those sequences might be treated as pseudogene rDNA alleles, or as highly divergent paralogues alleles (sensu Baldwin, 1995), or as class II paralogues (sensu Buckler et al., 1997) or deep paralogues (sensu Bailey et al., 2003). Furthermore, we report three additional ITS sequences from Manens' original data set as pseudogene rDNA alleles, i.e., *I. brasiliensis* AJ492661, *I. integerrima* AJ492664, and *I. pseudobuxus* AJ492660. In contrast, *I. microdonta* AJ492665 and *I. theezans* AJ492666 might represent alleles within different active ribosomal arrays. Our phylogenetic analysis including divergent ITS alleles, confirmed our results based on examining factors correlated to functionality. Different laboratory methods were employed in Manen's work and ours, concerning mainly the amplification primers (universal primer plus specific primer used herein, and both universal primers in Manen's) and the sequencing (direct sequencing of double stranded DNA was employed herein vs. cloning by Manen et al., 2002). These might account for the differential recovery of pseudogenes.

Our results support the idea that phylogeneticists using nuclear ITS sequences should be aware of the potential risks of downloading sequences from GenBank and using them for phylogenetic analysis without prior corroboration of their orthologous nature. As stated by Mayol and Roselló (2001), without a detailed inspection of some basic features, errors in evolutionary inferences could be easily overlooked. Incorporating ITS paralogues can distort the phylogenetic signal (Buckler et al., 1997) yielding spurious clades.

Phylogenetic inferences and sensitivity analyses—Direct optimization of DNA sequence data yields improved testing of phylogenetic hypotheses because provisional homologies are optimized for each cladogram individually, not a priori and universally as with the static homologies of multiple alignments (Wheeler, 2003). The sensitivity analysis as proposed by Wheeler (1995), results in the knowledge of nodal stability, showing how an alignment parameter affects the outcome tree topology (Giribet, 2003). Stability, logically different from support, can only be defined by reference to some factors (Goloboff et al., 2003). A group found in all the parameter space

may nonetheless be poorly supported, and at the same time, a group with high “support” may be unstable to changes in some parameters (Goloboff et al., 2003). A drawback of sensitivity analysis, is that the number of parameters to be explored is unbounded, and thus new or more gap : change cost ratios could always be added (Giribet and Wheeler, 1999). Trying to generalize an evolutionary model for all taxa, all loci, and all regions may be simplistic; however, evaluating two general parameters (gap/change ratio and Tv/Ts) is a starting point (Maxmen et al., 2003).

In this group of plants, analytical conditions rendered different hypotheses of relationships between species, indicating that the alignment parameters did influence the phylogenetic inference. Notwithstanding, for the ITS sequence data alone and in combination with AFLP data, the sensitivity analysis showed that some clades were stable to the variation in analytical parameters and that those clades mainly agreed with geographical origin of the taxa involved. For instance, the clade composed solely by southern South American taxa (the sSA clade) namely, *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, *I. microdonta*, *I. pseudobuxus*, *I. taubertiana*, and *I. theezans* was stable to parameter fluctuations (i.e., robust) in both analyses of ITS data alone and combined (Fig. 19). It is noteworthy that neither the sequences of *I. microdonta* nor those of *I. theezans* appeared monophyletic, which could be caused by ancestral polymorphism, introgression, or sampling errors. Another stable clade consisted of East Asian taxa *I. beecheyi*, *I. integra*, *I. kiusiana*, *I. mertensii*, and *I. percoriacea*. In addition, both stable clades were highly supported by the data in all conditions. On the other hand, East Asian clade formed by *I. cornuta*, *I. dimorphophylla*, and *I. maximowicziana* and the clade concerning North American taxa *I. collina*, *I. decidua*, and *I. vomitoria* were robust to fluctuations but showed variable support values. Many examples exist in the literature showing cases like the ones presented here, decoupling support from stability in both directions (Giribet, 2003).

Results presented here contradict those of Manen et al. (2002). A possible explanation for the incongruence of both studies would be based on the use of several highly divergent ITS sequences by Manen’s team and on the analytical methodology employed. In this last respect, in the work of Manen et al. (2002), sequences were aligned manually, gaps were treated as missing data and indels were coded as present or absent, but without an explicit criterion for doing so. Manual alignments have been criticized for the lack of objectivity, lack of repeatability, and lack of criterion for exclusion data (e.g., Gatesy et al., 1993; Giribet and Wheeler, 1999). In our study, the analysis proceeds directly from sequence data to phylogenetic reconstruction treating gaps as relevant evolutionary transformations (Giribet and Wheeler, 1999).

Herein, majority rule consensus trees were presented because the lack of an appropriate external criterion (i.e., character congruence, as suggested by Wheeler, 1995 and Giribet and Wheeler, 1999) prevents us from choosing among competing phylogenetic hypotheses. Taxonomic congruence is probably not a good external criterion when groups are defined arbitrarily (Giribet and Wheeler, 1999).

The AFLP topology (Fig. 2) and the majority rule consensus cladograms of ITS data alone and combined data (Fig. 19) were remarkably congruent in that they demonstrated the existence of a group formed by *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, and *I. theezans* and that *I. dumosa*

and *I. paraguariensis* were distantly related to the former species. These groupings were found to agree with the phenological data obtained by Coelho and Mariath (1996). Two main strategies of leaf longevity were distinguished by these authors, although some taxa were not included: subdeciduous (for *I. brevicuspis*, *I. integerrima*, *I. theezans*, *I. microdonta*, and *I. taubertiana*) and perennifoliar (for *I. dumosa* and *I. paraguariensis*).

Apart from performing separate analyses of ITS sequence data and AFLPs, we carried out a combined phylogenetic analysis. We found that robust clades obtained with ITS data alone were not affected by the addition of AFLP data. Only sSA clade exhibited any changes, such as a better resolution of the relationship between *I. brasiliensis*, *I. integerrima*, and *I. theezans*. Loesener (1901, 1908, 1942) and Baas (1975), based on shared morphological characters (for instance, inflorescences in fasciculate corymbs, flower and fruit of large sizes, coriaceous limbs with entire or pauci- serrate leaf margins), proposed a close relationship among these taxa. Other clades showed in the combined analysis an increment in their percentage of formation throughout the analytical space. However, no major differences in support values were observed between sequence data alone and the combined analysis.

In spite of their taxonomic classification, *I. pseudobuxus* and *I. taubertiana* failed to form a clade, suggesting that morphological features shared by these taxa (the occurrence of long and slender flower peduncles; cfr. Edwin and Reitz, 1967) might have been developed by convergence. Similarly, *I. microdonta* and *I. brevicuspis* did not group in accordance with their taxonomic disposition, and thus morphological characters that distinguish these taxa from the rest could have arisen by convergence.

Furthermore, our analyses revealed that changes in alignment parameters resulted in alternative taxa being placed as sister to the sSA clade and that Asian taxa were involved in the majority of analytical conditions. The placement of *I. paraguariensis* was likewise highly sensitive to alignment parameter variation, appearing either related to the *I. dumosa*-*I. crenata*-*I. mutchagara* group, or as a third independent sSA lineage (Figs. 7, 8, 10, 11, 14, 15, 17, and 18). With regard to *I. dumosa*, our results conflict with the work of Manen et al. (2002) in which *I. dumosa* appeared closely related to *I. argentina*, with both nuclear and plastid sequence data. We found no evidence of such a relationship, neither with AFLP data nor with ITS data. Instead, our results indicate a close relationship of *I. argentina* specimens to the remaining taxa of the sSA clade. It is worth mentioning that Manen’s sequence of *I. argentina* AJ492655 appeared herein, robustly related and highly supported by the data to North American *I. cassine* and *I. opaca*. The fact that both ITS sequences of *I. argentina* failed to form a monophyletic clade could be explained by the coexistence within this tetraploid species of many types of active ITS copies due to an incomplete homogenization or to a recent formation of the polyploid. Alternatively, the possibility of mislabeled samples cannot be discarded. It would be interesting to reanalyze all the data set used by Manen et al. (2002), including plastid sequence data, through optimization and sensitivity analyses, in order to evaluate the stability of the phylogenetic hypotheses proposed. Present results support the idea that current classification of *I. argentina*, closely allied to *I. paraguariensis* (Baas, 1975), may not be justified and that a re-classification should be accomplished.

Our phylogenetic analysis provides no examples of complete agreement between stable clades and the taxonomical treatment of Loesener (1891, 1901, 1908, 1942). Other studies, such as those of Cuénoud et al. (2000) and Manen et al. (2002), indicated an inconsistency of Loesener's infrageneric system with sequence data. This inconsistency was also noted with Galle's system (not shown), which basically summarized previous taxonomic works, but using contradictory criteria. The phylogenetic study of Setoguchi and Watanabe (2000) that comprises solely Asian *Ilex* taxa followed the taxonomic treatment of Yamazaki (1989). In our opinion, a discrepancy exists between Yamazaki's treatment and sequence data presented by those authors. Contradictions between morphological and molecular data are somehow expected because, as explained by Systema (1990), reliance on overall morphological similarity as an approximation of phylogenetic relationships rather than on shared derived character states can generate misleading estimates.

Our results draw attention to the need of a thorough taxonomic revision of *Ilex* that would be based on a simultaneous cladistic analysis of all morphological and molecular characters available, without disregarding the influence of analytical parameters on phylogenetic reconstruction.

LITERATURE CITED

- ALBACH, D. C., P. S. SOLTIS, D. E. SOLTIS, AND R. G. OLMSTEAD. 2001. Phylogenetic analysis of asterids based on sequences of four genes. *Annals of the Missouri Botanical Garden* 88: 163–212.
- ALICE, L. A., AND C. S. CAMPBELL. 1999. Phylogeny of *Rubus* (Rosaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. *American Journal of Botany* 86: 81–97.
- ALVAREZ, I., AND J. F. WENDEL. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417–434.
- BAAS, P. 1975. Vegetative anatomy and the affinities of Aquifoliaceae, *Sphenostemon*, *Phelline* and *Oncothecca*. *Blumea* 22: 311–407.
- BAILEY, C. D., T. G. CARR, S. A. HARRIS, AND C. E. HUGHES. 2003. Characterization of angiosperm nrDNA polymorphism, paralogy and pseudogenes. *Molecular Phylogenetics and Evolution* 29: 435–455.
- BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, AND M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247–277.
- BARRAL, G., L. POGGIO, AND G. C. GIBERTI. 1995. Chromosome numbers and DNA content from *Ilex argentina* (Aquifoliaceae). *Boletín de la Sociedad Argentina de Botánica* 30: 243–248.
- BASSAM, B. J., G. CAETANO-ANOLLÉS, AND P. M. GRESHOFF. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* 196: 81–84.
- BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.
- BREMER, B., K. BREMER, N. HEIDARI, P. ERIKSON, R. G. OLMSTEAD, A. A. ANDERSBERG, M. KÄLLERSJÖ, AND E. BARKHORDARIAN. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Molecular Phylogenetics and Evolution* 24: 274–301.
- BUCKLER, E. S., AND T. P. HOLTSFORD. 1996a. *Zea* systematics: ribosomal ITS evidence. *Molecular Biology and Evolution* 13: 612–622.
- BUCKLER, E. S., AND T. P. HOLTSFORD. 1996b. *Zea* ribosomal repeat evolution and substitution patterns. *Molecular Biology and Evolution* 13: 623–632.
- BUCKLER, E. S., A. IPPOLITO, AND T. P. HOLTSFORD. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* 145: 821–832.
- CAMPBELL, C. S., M. F. WOJCIECHOWSKI, B. G. BALDWIN, L. A. ALICE, AND M. J. DONOGHUE. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). *Molecular Biology and Evolution* 14: 81–90.
- CERVERA, M. T., J. A. CABEZAS, J. C. SANCHA, F. MARTÍNEZ DE TODA, AND J. M. MARTÍNEZ-ZAPATER. 1998. Application of AFLPs to the characterization of grapevine *Vitis vinifera* L., genetic resources. A case of study with accessions from Rioja (Spain). *Theoretical and Applied Genetics* 97: 51–59.
- COELHO, G. C., AND J. E. A. MARIATH. 1996. Inflorescence morphology of *Ilex* L. (Aquifoliaceae) species from Rio Grande do Sul, Brazil. *Feddes Repertorium* 107: 19–30.
- CUÉNOUD, P., M. A. DEL PERO MARTÍNEZ, P.-A. LOIZEAU, R. SPICHTER, S. ANDREWS, AND J.-F. MANEN. 2000. Molecular phylogeny and biogeography of the genus *Ilex* L. (Aquifoliaceae). *Annals of Botany, London* 85: 111–122.
- DE LAET, J., AND W. C. WHEELER. 2003. Poy version 3.0.11 (Wheeler, Gladstein and DeLaet, 6 May 2003). Computer software and documentation distributed by the authors at website, <http://research.amnh.org/scicomp/projects/poy.php>.
- DOWNIE, S. R., D. S. KATZ-DOWNIE, AND K. SPALIK. 2000. A phylogeny of Apiaceae tribe Scandiceae: evidence from molecular ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany* 87: 76–95.
- DOYLE, J. J., AND J. I. DAVIES. 1998. Homology in molecular phylogenetics: a parsimony perspective. In D. E. Soltis, P. S. Soltis, and J. J. Doyle [eds.], *Molecular systematics of plants II. DNA sequencing*, 101–103. Kluwer, Boston, Massachusetts, USA.
- EDWIN, G., AND R. REITZ. 1967. Aquifoliaceae. In R. Reitz [ed.], *Flora illustrada catarinense*. Tipografia e Livraria Blumenauense S. A. Itajaí, Brazil.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 39: 783–791.
- FILIP, R., P. LÓPEZ, G. C. GIBERTI, J. COUSSIO, AND G. FERRARO. 2001. Phenolic compounds in seven South American *Ilex* species. *Fitoterapia* 72: 774–778.
- GALLE, F. C. 1997. *Hollies: the genus Ilex*. Timber Press, Portland, Oregon, USA.
- GATESY, J., R. DESALLE, AND W. C. WHEELER. 1993. Alignment of ambiguous nucleotide sites and the exclusion of systematic data. *Molecular Phylogenetics and Evolution* 2: 152–157.
- GAUER, L., AND S. CAVALLI-MOLINA. 2000. Genetic variation in natural populations of maté (*Ilex paraguariensis* A. St. Hil., Aquifoliaceae) using RAPD markers. *Heredity* 84: 647–654.
- GIBERTI, G. C. 1994a. Aquifoliaceae. In R. Spichiger and L. Ramella (Dirs.), *Flora del Paraguay*, vol. 24, Editions des Conservatoire et Jardin botaniques de la Ville de Genève, Missouri Botanical Garden, Héliographia S A, Geneva, Switzerland.
- GIBERTI, G. C. 1994b. Aquifoliaceae. In A. T. Hunziker and A. M. Anton [eds.], *Flora fanerogámica de Argentina*, vol. 1, 1–8, Programa PROFLORA (CONICET), Museo Botánico, IMBIV, Córdoba, Estudio Sigma S.R.L., Buenos Aires, Argentina.
- GIBERTI, G. C. 1998. Hallazgo de *Ilex brasiliensis* (Aquifoliaceae) en la Argentina. *Boletín de la Sociedad Argentina de Botánica* 33: 137–140.
- GIRIBET, G. 2001. Exploring the behavior of POY, a program for direct optimization of molecular data. *Cladistics* 17: S60–S70.
- GIRIBET, G. 2003. Stability in phylogenetic formulations, and its relationship to nodal support. *Systematic Biology* 52: 554–564.
- GIRIBET, G., AND C. RIBERA. 2000. A review of arthropod phylogeny: new data based on ribosomal DNA sequences and direct character optimization. *Cladistics* 16: 204–231.
- GIRIBET, G., AND W. C. WHEELER. 1999. On gaps. *Molecular Phylogenetics and Evolution* 13: 132–143.
- GLADSTEIN, G., AND W. C. WHEELER. 1997. POY: the optimization of alignment characters. American Museum of Natural History, New York, New York, USA.
- GOLOBOFF, P. A., J. S. FARRIS, M. KÄLLERSJÖ, B. OXELMAN, M. J. RAMÍREZ, AND C. A. SZUMIK. 2003. Improvements to resampling measures of group support. *Cladistics* 19: 324–332.
- GRAY, A. 1856. *Manual of the botany of the northern United States*, 2nd ed., including Virginia, Kentucky, and all East of the Mississippi. G. P. Putnam and Co., New York, New York, USA.
- HARTMANN, S., J. D. NASON, AND D. BHATTACHARYA. 2001. Extensive ribosomal DNA genic variation in the columnar cactus *Lophocereus*. *Journal of Molecular Evolution* 53: 124–134.
- HALL, T. A. 1999. BioEdit: a user-friendly biological sequence alignment

- editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
- HOOKER, J. D. 1862. Ilicineae. In G. Bentham and J. D. Hooker [eds.], *Genera plantarum*, vol. 1, 355–357. A Black, London, UK.
- HU, S.-Y. 1949. The genus *Ilex* in China. *Journal of the Arnold Arboretum, Harvard University* 30: 233–344, 348–387.
- HU, S.-Y. 1950. The genus *Ilex* in China. *Journal of the Arnold Arboretum, Harvard University* 31: 39–80, 214–240, 241–263.
- KARDOLUS, J. P., H. J. VAN ECK, AND R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). *Plant Systematics and Evolution* 210: 87–103.
- KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111–120.
- KOOPMAN, W. J. M., M. J. ZEVENBERGER, AND R. G. VAN DEN BERG. 2001. Species relationships in *Lactuca* s. l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. *American Journal of Botany* 88: 1881–1887.
- KRÜSSMANN, G. 1962. *Handbuch der Laubgehölze in zwei Bänden*, band II. Paul Parey, Berlin, Germany.
- LIA, V. V., V. A. CONFALONIERI, C. I. COMAS, AND J. H. HUNZIKER. 2001. Molecular phylogeny of *Larrea* and its allies (Zygophyllaceae): reticulate evolution and the probable time of creosote bush arrival to North America. *Molecular Phylogenetics and Evolution* 21: 309–320.
- LILLO, M. 1911. Descripción de plantas nuevas pertenecientes a la flora Argentina. *Anales de la Sociedad Científica Argentina* 72: 171–175.
- LINNAEUS, C. 1753. *Species plantarum*, vol. 1. Impensis Laurentii Salvii, Stockholm, Sweden.
- LIU, J.-S., AND C. L. SCHARDL. 1994. A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. *Plant Molecular Biology* 26: 775–778.
- LOBREAU-CALLEN, D. 1975. Les variations dimensionnelles du pollen du genre *Ilex* (Aquifoliaceae) et leur rapport avec le climat. *Bulletin de la société botanique de France. Collection Palynologie* 122: 179–199.
- LOESENER, T. 1891. Vorstudien zu einer Monographie der Aquifoliaceae. *Verhandlungen den Botanischen Vereins für die Provinz Brandenburg und die angrenzenden Ländern* 33: 1–45.
- LOESENER, T. 1901. Monographia aquifoliacearum. I. *Nova Acta Academiae Caesareae Leopoldinae—Carolinae Germanicae Naturae Curiosorum* 78.
- LOESENER, T. 1908. Monographia aquifoliacearum. II. *Nova Acta Academiae Caesareae Leopoldinae—Carolinae Germanicae Naturae Curiosorum* 89.
- LOESENER, T. 1942. Aquifoliaceae. In H. Harms and J. Matfeld [eds.], *Die Natürlichen Pflanzenfamilien* 2nd ed., 36–86, Wilhelm Engelmann, Leipzig, Germany.
- MACE, E. S., C. G. GEBHARDT, AND R. N. LESTER. 1999. AFLP analysis of genetic relationship in the tribe Datureae (Solanaceae). *Theoretical and Applied Genetics* 99: 634–641.
- MANEN, J.-F., M. C. 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 Systematics and Evolution* 235: 79–98.
- MARTIN, H. 1977. The history of *Ilex* (Aquifoliaceae) with special reference to Australia: evidence from pollen. *Australian Journal of Botany* 25: 655–673.
- MAYOL, M., AND J. A. ROSELLÓ. 2001. Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. *Molecular Phylogenetics and Evolution* 19: 167–176.
- MAXMEN, A. B., F. K. BURNETT, E. B. CUTLER, AND G. GIRIBET. 2003. Evolutionary relationships within the protostome phylum Sipuncula: a molecular analysis of ribosomal genes and histone H3 sequence data. *Molecular Phylogenetics and Evolution* 27: 489–503.
- MORRISON, D. A., AND J. T. ELLIS. 1997. Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18 S rDNA of Apicomplexa. *Molecular Biology and Evolution* 14: 428–441.
- MUELLER, U. G., AND L. WOLFENBARGER. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* 14: 389–394.
- NEGI, M. S., A. SINGH, AND M. LAKSHMIKUMARAN. 2000. Genetic variation and relationship among and within *Withania* species as revealed by AFLP markers. *Genome* 43: 975–980.
- NEI, M., AND W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: S269–S273.
- OGDEN, H. T., AND M. F. WHITING. 2003. The problem with “the Paleoptera Problem”: sense and sensitivity. *Cladistics* 19: 432–442.
- PAGE, R. D. M. 1998. TreeView 1.5.2. Computer software and documentation distributed by the author at website, <http://taxonomy.zoology.gla.ac.uk/rod/rod.html>.
- POWELL, W., M. MORGANTE, C. ANDRE, M. HANGLEY, J. VOGEL, S. TINGEY, AND A. RAFALSKI. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellites) markers for germplasm analysis. *Molecular Breeding* 2: 225–238.
- POWELL, M., V. SAVOLAINEN, P. CUÉNOUD, J.-F. MANEN, AND S. ANDREWS. 2000. The mountain holly (*Nemopanthus mucronatus*, Aquifoliaceae) revisited with molecular data. *Kew Bulletin* 55: 341–347.
- REHDER, A. 1927. Manual of cultivated trees and shrubs hardy in North America exclusive of the subtropical and warmer temperate regions. Macmillan, New York, New York, USA.
- REISSEK, S. 1861. Celastraceae, Ilicineae et Rhamneae. In C. F. P. von Martius [ed.], *Flora brasiliensis* 11, fasc. 28, Lipsiae ap. F. Fleischer, Leipzig, Germany.
- REN, N., AND M. P. TIMKO. 2000. AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. *Genome* 44: 559–571.
- ROBINSON, J. P., AND S. A. HARRIS. 1999. Amplified fragment length polymorphism and microsatellites: a phylogenetics perspective. In E. M. Gillet [ed.], Which DNA marker for which purpose? Available online <http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm>.
- SAVOLAINEN, V., M. W. CHASE, S. B. HOOT, C. M. MORTON, D. E. SOLTIS, C. BAYER, M. F. FAY, A. Y. DE BRUIJN, S. SULLIVAN, AND Y.-L. QIU. 2000. Phylogenetics of flowering plants based on combined analysis of plastid *atpB* and *rbcL* gene sequences. *Systematic Biology* 49: 306–362.
- SETOGUCHI, H., AND I. WATANABE. 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.
- STATSOFT, INC. 1999. STATISTICA version 5.5 A for Windows [computer program manual]. StatSoft, Inc., Tulsa, Oklahoma, USA. Also available online <http://www.statsoft.com/>
- SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods), Version 4b10. Sinauer Associates, Sunderland, Massachusetts, USA.
- SYSTMA, K. J. 1990. DNA and morphology: inference of plant phylogeny. *Trends in Ecology and Evolution* 5: 104–110.
- VILGALYS, R., AND M. HESTER. 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. *Journal of Bacteriology* 172: 4238–4246.
- VOS, P., R. HOGERS, M. BLEEKER, M. REIJNS, T. VAN DE LEE, M. HORNES, A. FRITJERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.
- WHEELER, W. C. 1993. The triangle inequality and character analysis. *Molecular Biology and Evolution* 10: 707–712.
- WHEELER, W. C. 1995. Sequence alignment, parameter sensitivity and the phylogenetic analysis of molecular data. *Systematic Biology* 44: 321–331.
- WHEELER, W. C. 1996. Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* 12: 1–9.
- WHEELER, W. C. 2002. Optimization alignment: down, up, error and improvements. In R. DeSalle, G. Giribet, and W. C. Wheeler [eds.], *Techniques in molecular systematics and evolution*, 55–59, Birkhäuser, Basel, Switzerland.
- WHEELER, W. C. 2003. Implied alignment: a synapomorphy-based multiple-sequence alignment method and its use in cladogram search. *Cladistics* 19: 261–268.
- WHEELER, W. C., AND D. GLADSTEIN. 2000. POY. Program and documentation, available at <ftp://amnh.org/pub/molecular>.
- WHITE, T. J., T. BRUNS, S. LEE, AND J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. Innis, D. Gelfand, J. Sininsky, and T. White [eds.], *PCR protocols. A guide to methods and applications*, 315–322. Academic Press, San Diego, California, USA.
- YAMAZAKI, T. 1989. Aquifoliaceae. In Y. Sotabe, H. Hara, S. Watari, and T. Tominari [eds.], *Wild flowers of Japan*, vol. 2, Woody plants, 26–32, Heibonsha, Tokyo (in Japanese).
- ZHU, J., M. D. GALE, S. QUARRIE, M. T. JACKSON, AND G. J. BYAN. 1998. AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics* 96: 602–611.
- ZUKER, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 31: 3406–3414.