American Journal of Botany 90(8): 1229–1238. 2003.

A FIRST MOLECULAR PHYLOGENETIC ANALYSIS OF *PASSIFLORA* (PASSIFLORACEAE)[1]

VALÉRIA C. MUSCHNER,[2] ALINE P. LORENZ,[2] ARMANDO C. CERVI,[3] SANDRO L. BONATTO,[4] TATIANA T. SOUZA-CHIES,[5] FRANCISCO M. SALZANO,[2] AND LORETA B. FREITAS[2,6]

[2]Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil; [3]Departamento de Botânica, Universidade Federal do Paraná, Caixa Postal 19031, 81531-970 Curitiba, PR, Brazil; [4]Centro de Biologia Genômica e Molecular, Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Ipiranga 6681, 90610-001 Porto Alegre, RS, Brazil; and [5]Departamento de Botânica, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Prédio 43433, 90509-900 Porto Alegre, RS, Brazil

Passiflora, a genus with more than 400 species, exhibits a high diversity of floral and vegetative structures and a complex taxonomy, which includes 23 subgenera and many sections and series. To better understand *Passiflora*'s variability and interspecific relationships, the phylogeny of 61 species, classified in 11 of 23 suggested subgenera, was investigated. Three molecular markers were used, the nuclear ribosomal internal transcribed spacers (nrITS), the plastid *trnL-trnF* spacer regions (~1000 bp), and the *rps4* plastid gene (~570 bp). Three major clades were highly supported, independent of the marker and phylogenetic method used; one included the subgenera *Distephana*, *Dysosmia*, *Dysosmioides*, *Passiflora*, and *Tacsonioides*, a second, the subgenera *Adopogyne*, *Decaloba*, *Murucuja*, and *Pseudomurucuja*, and a third, the subgenus *Astrophea*. We call these the *Passiflora*, *Decaloba*, and *Astrophea* clades, respectively. The position of subgenus *Deidamioides* is undefined. The monophyly of *Passiflora* could not be statistically corroborated, and the relationships among the major clades and of these clades with the related genera remain unresolved. Our results indicate that a reevaluation of the monophyly of *Passiflora* and its infrageneric classification is necessary.

Key words: ITS; *Passiflora*; Passifloraceae; phylogenetic analysis; *rps4*; *trnL-trnF*.

Following Escobar (1988), Passifloraceae are divided into two tribes: Paropsieae (with six genera, all Old World species in Africa and Madagascar) and Passifloreae (with 14 genera, five of them present in the New World and nine in the Old World). Judd et al. (1999) suggested that the monophyly of Passifloraceae is supported most strongly by the presence of a corona in the flowers. The Paraopsieae, which contains shrubs and trees that lack tendrils, probably represent a paraphyletic basal complex within the family. Passifloreae, in contrast, are clearly monophyletic, as evidenced by their vine habit, axillary tendrils, and specialized flowers. The genus *Passiflora*, which includes the largest number of species described thus far (about 400 for the New World and 20 for Asia; Cervi, 1997), is included among the Passifloreae. The taxonomy of *Passiflora* is based on several floral and vegetative structures, leading to a complex taxonomic subdivision in subgenera, sections, and series. As a matter of fact, Killip (1938) and MacDougal (1994) asserted that among angiosperms no other group presents such diversity in leaf form. Its flowers also vary widely in size and color, with the corona and perianth diversely oriented and developed, all of which may have arisen as a result of coevolutionary relationships with insect pollinators (MacDougal, 1994).

According to Killip (1938), the genus should be subdivided into 22 subgenera and the *Decaloba* and *Astrophea* subgenera into 13 sections. Escobar (1989) described one additional subgenus. Killip's (1938) classification was entirely based on morphological characters, especially those of the floral structure. Only two other studies deal with the infrageneric relationships in *Passiflora* (Sánchez et al., 1999; de Melo et al., 2001), but none employed modern phylogenetic methods, DNA sequence data, or suggested any infrageneric grouping.

In the present investigation we analyzed the relationships of 61 species of *Passiflora*, formally classified in 11 subgenera, and representatives of four other genera, using two noncoding DNA segments, the nuclear ribosomal internal transcribed spacers (ITS), and the plastid *trnL-trnF* spacers. These regions have been widely used in plant phylogenetics because of their high rate of nucleotide substitutions (Taberlet et al., 1991; Baldwin, 1992; Baldwin et al., 1995; Rauscher, 2002) and their power in elucidating infrageneric relationships (Gielly and Taberlet, 1996; Bakker et al., 1998; Molvray et al., 1999; Nishikawa et al., 1999; Hardig et al., 2000). The *rps4* gene (encoding protein 4 of the small plastid ribosomal subunit) was also studied in a smaller number of species to broadly access the major phylogenetic relationships within *Passiflora*. This gene was chosen mainly because of its slower evolutionary rate when compared with the other two regions studied here and its previously successful use in the phylogenetic reconstruction of Poaceae (Nadot et al., 1994) and Iridaceae (Souza-Chies et al., 1997). The data were evaluated using several phylogenetic reconstruction methods. Specific objectives were (a) to examine the nature of the variation in these two differently positioned spacer regions and in a coding region, assessing

[1] Manuscript received 12 September 2002; revision accepted 28 March 2003.

The authors thank Mark Chase (Royal Botanic Gardens, Kew, UK), Maurizio Vecchia, Marcelo S. Guerra-Filho, Natoniel Franklin de Melo, Cláudio Mondin, Teonildes Nunes, and Marcelo C. Dornelas for specimens; Cassio van den Berg for critical review of the manuscript; Valério Pilar for computational facilities; and an anonymous reviewer for helpful comments and suggestions on an earlier draft of this manuscript. Financial support was provided by Programa de Apoio a Núcleos de Excelência (PRONEX), Financiadora de Estudos e Projetos (FINEP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), and Pró-Reitoria de Pesquisa da Universidade Federal do Rio Grande do Sul (PROPESQ-UFRGS).

[6] Author for correspondence (e-mail: loreta.freitas@ufrgs.br).

their phylogenetic utility in *Passiflora*; (b) to investigate the monophyly of the genus; and (c) to clarify the infrageneric phylogenetic relationships.

MATERIALS AND METHODS

Plants studied and laboratory methods—The 61 species representing 11 subgenera investigated here are listed in Appendix 1 (see Supplementary Data accompanying the online version of this article), together with representatives from four other genera of Passifloraceae (*Adenia*, *Mitostemma*, *Tetrastylis*, *Paropsia*) utilized as outgroups. This sampling includes all species of Passifloraceae from which we could obtain suitable material to extract DNA, containing taxa from a wide distributional range in South and Central America. Among the outgroups, representatives of the two tribes of Passifloraceae are considered.

DNA was extracted from fresh leaves (collected directly from nature or derived from seeds grown in the laboratory) or from herbarium material, using the cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1987). Amplification was done by polymerase chain reaction (PCR), using the primers described by Desfeux and Lejeune (1996) for ITS and the primers *e* and *f* of Taberlet et al. (1991) for *trnL-trnF*. The *rps4* gene was amplified with primers rps5 and trnS described by Souza-Chies et al. (1997). The amplified material was purified with shrimp alkaline phosphatase and exonuclease I (Amersham Biosciences, Piscataway, New Jersey, USA), and the two strands were directly sequenced, using the internal primers ITS3 and ITS3-reverse (Desfeux and Lejeune, 1996) for ITS. Sequencing was done using the BigDye Terminator Sequencing Kit (PE Applied Biosystems, Foster City, Michigan, USA) on an ABI Prism 310 (PE Applied Biosystems). Sequences were deposited with GENBANK (see Appendix 1 in the Supplementary Data accompanying the online version of this article).

Data analysis—Sequences were aligned using the ClustalX 1.81 program (Thompson et al., 1994, 2001) and manually refined. Phylogenetic analyses were performed in PAUP*, version 4.0b10 (Swofford, 1998). All the analyses described later were done with each spacer considered as a different data set (ITS or *trnL-trnF*) and also with both markers combined in a single alignment. ITS1 and ITS2 were considered together. Because of technical problems, we could not obtain usable sequences for some species for some markers, so that only 55 species were studied for ITS and 60 for *trnL-trnF*. For *rps4*, we sequenced material from 32 species, in order to obtain representatives of the major clades of *Passiflora*, which were compared with three of the four outgroups.

Equally weighted parsimony (maximum parsimony [MP]) analyses were performed by a heuristic search with tree bisection-reconnection (TBR) branch swapping, the MULPARS option, and 1000 random-addition replicates. Bootstrap statistical support (Felsenstein, 1985) was carried out with 1000 replications of heuristic search and simple taxon addition, with the all trees saved option. Topological congruence between the trees obtained by the different data sets was evaluated with Templeton's test (1983), and heterogeneity among markers was assessed using the partition homogeneity test or incongruence length difference test (IDL) developed by Farris et al. (1994) with 1000 iterations. The g1 statistic (Hillis, 1991) of skewed tree-length distribution was calculated from 10 000 random trees, to measure the phylogenetic information content of the two spacer regions independently and for the combined data.

The appropriate model of nucleotide substitution for maximum-likelihood (ML) analysis was determined using the MODELTEST 3.06 program (Posada and Crandall, 1998). This procedure implements a hierarchical likelihood-ratio test to determine the model that best fits the data. For ML tree estimation, heuristic searches with as-is, TBR branch swapping, and MULPARS option, were conducted with PAUP*. Because obtaining support estimates for the ML trees' branches by bootstrap analysis is computationally too intensive with a data set of this size, we conducted a bootstrap analysis using the neighbor-joining method under a maximum-likelihood model, with parameters settings estimated by the MODELTEST, as described by Xiang et al. (2002). Topology comparisons between the trees obtained with the two spacers were done with the KH (Kishino and Hasegawa, 1989) and SH (Shimodaira and Hasegawa, 1999) tests.

The new MetaPIGA 1.02 software was used for phylogenetic inference under the ML criterion using a metapopulation genetic algorithm (Lemmon and Milinkovitch, 2002). This method was developed to supply better and faster estimates of ML trees with large data sets. The program was run using the following relevant settings: HKY85 (Hasegawa et al., 1985) nucleotide substitution model, four populations, and 10 replications. Lemmon and Milinkovitch (2002) suggest that four population metaPIGA runs seem to give the best compromise between computing speed and accuracy of topology inference. All other settings were used as suggested in the program manual (Lemmon and Milinkovitch, 2002). The 50% majority-rule consensus tree was constructed in PAUP*.

Bayesian analyses of the data were performed using MrBayes 2.01 (Huelsenbeck and Ronquist, 2001) to generate a posterior probability distribution using Markov chain Monte Carlo (MCMC) methods. No a priori assumptions about the topology of the tree were made and all searches were provided with a uniform prior. The models of DNA substitution used were those estimated by MODELTEST and used in the standard ML analyses. The MCMC processes were set so that four chains were run simultaneously for 500 000 generations, with trees being sampled every 100 generations for a total of 5000 trees. Burn-in, or the time for each parameter to reach a stationary state, was determined when visual inspection indicated that the log-likelihood values achieved an asymptote over a large number of generations. To calculate the posterior probability of each bipartition, a 50% majority-rule consensus tree was constructed from the remaining trees using PAUP*.

For the distance analyses, trees were constructed using the neighbor-joining method (NJ; Saitou and Nei, 1987) using proportional (p), Kimura two parameter, and logDet (Steel, 1994; Lockhart et al., 1994) distances. LogDet or paralinear distances were calculated to test the possible influence of the nucleotide composition difference in the phylogeny (Nei and Kumar, 2000). Reliability of the trees was tested using 1000 bootstrap replications (Hedges, 1992). Topology comparisons between the ITS and *trnL-trnF* data sets were performed using the KH and SH tests. Relative rate evaluations, based on the two-cluster test of Takezaki et al. (1995), were performed using the PHYLTEST program (Kumar, 1996). Average nucleotide diversities and their standard errors within each subgenus were calculated by the SendBS program (Takezaki, 2001) with the *p* distance option and 2000 bootstrap replications.

The results of the different inference phylogenetic methods within each data set were quite similar, so only selected results are presented here. (All other trees are presented in the Supplementary Data accompanying the online version of this article.) All trees were rooted with *Paropsia* because it is representative of Paropsieae and was shown to be more basal than *Adenia* by Chase et al. (2002).

A possible substitution saturation in the ITS region was investigated by plotting the ITS pairwise sequence *p* distances against the more conserved *trnL-trnF* pairwise distances. The minimum sizes of the flowers (using the descriptions of Killip [1938], MacDougal [1994], and Cervi [1997]) of different clades (see Results) were compared using the Wilcoxon-Mann-Whitney test to determine if their averages were significantly different.

RESULTS

Sequence variation—The ITS alignment is 586 bp in length (individual sequences ranged from 387 to 479 bp) (ITS1 is from position 1 to 353 and ITS2 from position 354 to the final) and is highly variable with 347 parsimony-informative sites. Because of the high divergence, the ITS alignment among the sequences of the different genera is ambiguous in several regions and many indels are present. Interestingly, all species of subgenera *Passiflora*, *Dysosmia*, *Dysosmioides*, *Tacsonioides*, and *Distephana* have sequences shorter than the other species, possibly from a large (about 50 bp) deletion around position 58. Comparison with other genera suggests that this deletion may be synapomorphic. The guanine-cytosine (GC) content is high in *Passiflora* (63%), but it is significantly lower in sub-

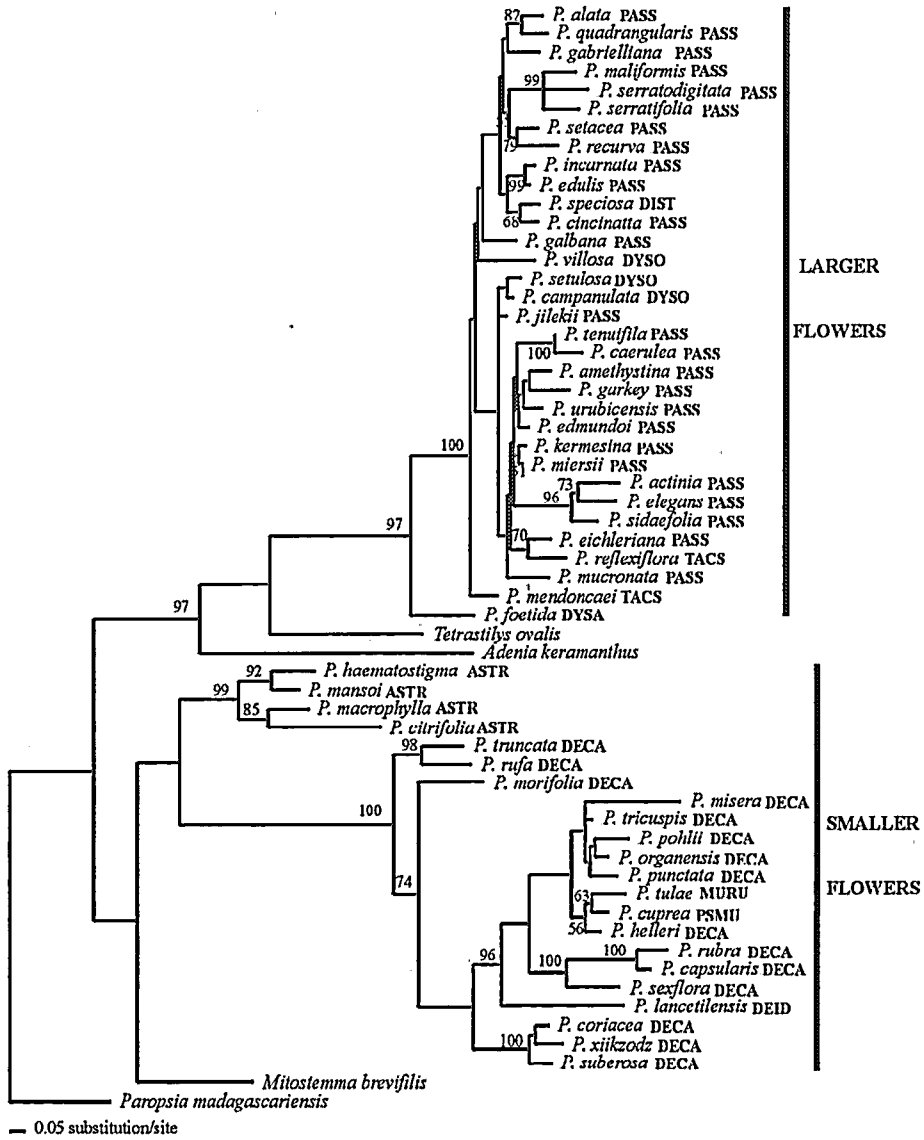

Fig. 1. Maximum-likelihood tree for the ITS spacer in *Passiflora* and outgroups. Numbers above branches are bootstrap support values (when higher than 50%) based on 1000 replicates. Abbreviations indicate the subgenera as follows: PASS = *Passiflora*; DIST = *Distephana*; TACS = *Tacsonioides*; DYSO = *Dysosmioides*; DYSA = *Dysosmia*; ASTR = *Astrophea*; DECA = *Decaloba*; MURU = *Murucuja*; DEID = *Deidamioides*; PSMU = *Pseudomurucuja*; ADOP = *Adopogyne*.

genera *Decaloba*, *Adopogyne*, *Murucuja*, *Pseudomurucuja*, and *Deidamioides* (53%), as compared to all the other subgenera (*Passiflora*, *Dysosmia*, *Dysosmioides*, *Distephana*, and *Tacsonioides*) (67%) ($\chi^2 = 34.8$; 1 df; $P < 0.0001$). The pairwise *p* distances, including the outgroups, ranged from 1.9% to 38%, with an average value of 20%.

The *trnL-trnF* intergenic region had a total aligned length of 416 bp (individual sequences ranged from 239 to 313 bp), 61 being parsimony-informative. The pairwise *p* distances, including the outgroups, ranged from 1.8% to 10%, with an average value of 4.4%. The GC content is low in this region (36%), and no marked differences were found among the subgenera. The alignment of the *rps4* gene is 576 bp in length (unaligned sequences ranged from 542 to 560 bp), 58 being parsimony-informative. The GC content is very low in this region (27%), and again, no marked differences were found among the subgenera. The pairwise *p* distances, including the outgroups, ranged from 0.2% to 8.1%, with an average value of 3.5%.

Phylogenetic analysis—The selected model for the maximum-likelihood analysis using MODELTEST was TrN+I+G (−lnL = 8227.1035) for ITS, HKY + G (−lnL = 2131.0359) for *trnL-trnF*, and TVM + G (−lnL = 2164.3538) for *rps4* (see Posada and Crandall [1998] for details of the models). The gamma shape parameters estimated for the three DNA regions were also different (ITS, 1.839; *trnL-trnF*, 0.5640; *rps4*, 0.7806). For the combined data of the spacer regions, the model selected was TrNef+I+G (−lnL = 10381.6836) with a gamma shape parameter of 1.078. Figures 1–3 present the phylogenetic trees obtained using the standard ML method on the ITS, *trnL-trnF*, and combined data sets, respectively. In the great majority of the trees we can observe three major clades within *Passiflora*, which are generally supported by high bootstrap values: one composed of all studied species of subgenera *Distephana*, *Dysosmia*, *Dysosmioides*, *Passiflora*, and *Tacsonioides*; another comprising all species of sugenera *Adopogyne*, *Decaloba*, *Murucuja*, and *Pseudomurucuja*; and a third composed of the four species of subgenus *Astrophea*. We

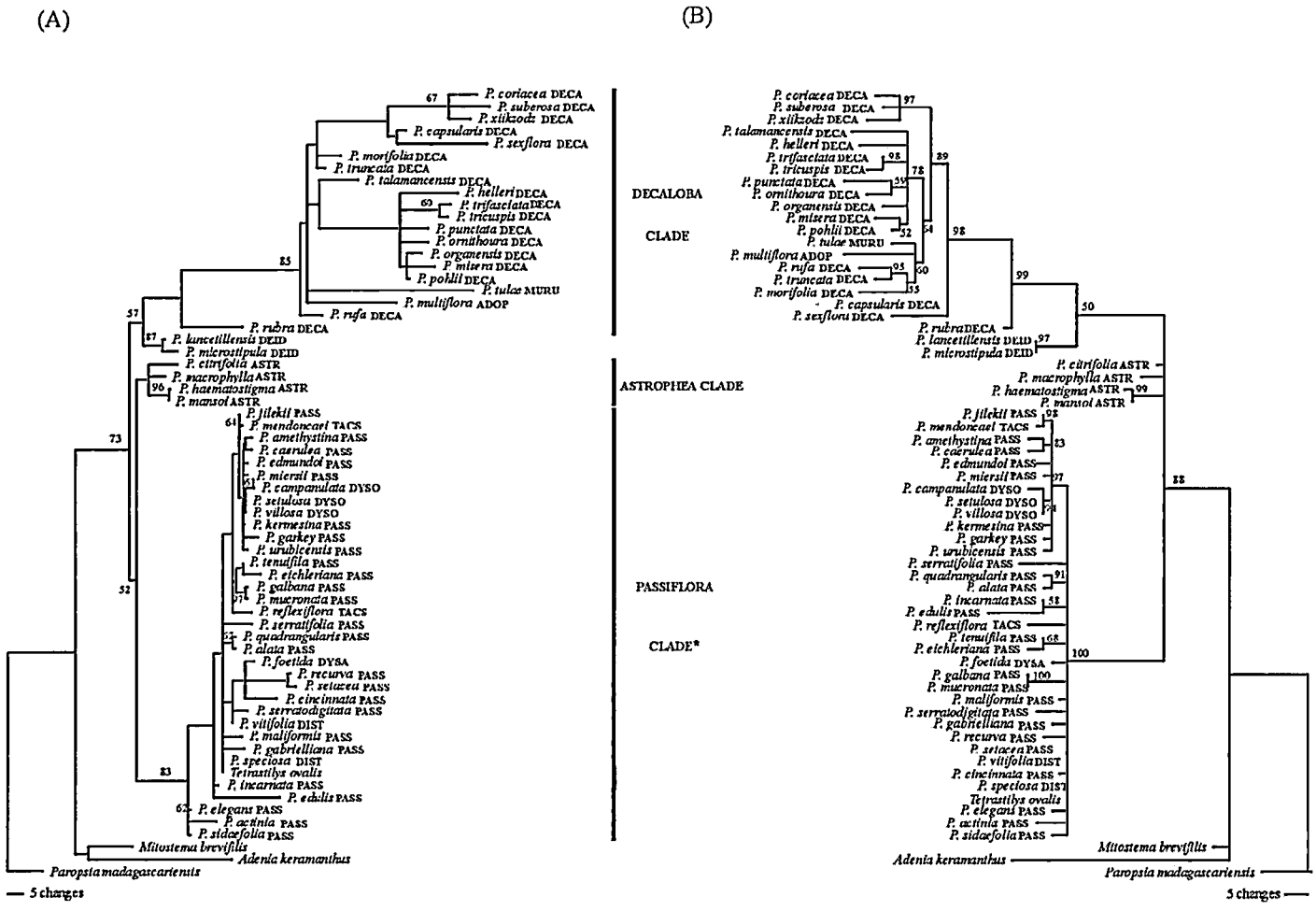

Fig. 2. Phylogenetic tree for the intergenic spacer *trnL-trnF* in *Passiflora* and outgroups. (A) Maximum-likelihood tree. Numbers above branches are bootstrap support values (when higher than 50%) based on 1000 replicates. (B) Bayesian analysis tree. Numbers above branches are posterior probabilities values (when higher than 50%). Abbreviations indicate the subgenera (see Fig. 1 for full names). * Except *Tetrastilys ovalis*.

call these the *Passiflora*, *Decaloba*, and *Astrophea* clades, respectively.

The trees obtained by the Bayesian approach and the metaPIGA program were strikingly similar among themselves and with the standard ML approach inferred by PAUP* (e.g., Fig. 2). The support values for most of the tree branches were consistently higher in the Bayesian and metapopulation genetic algorithm methods than the bootstrap values estimated in the standard ML method (Fig. 2). These results agree with the simulations performed by Wilcox et al. (2002) comparing the Bayesian posterior probabilities support values with the nonparametric bootstrap values and with the results of Lemmon and Milinkovitch (2002) on the metaGA approach. They concluded that the higher support values for the Bayesian and metaGA analyses are appropriate and provide much closer estimates of phylogenetic accuracy than the estimates given by corresponding bootstrap proportions (but see Suzuki et al. [2002] for a rebuttal).

In the maximum-parsimony analyses, the heuristic search of the ITS data resulted in 72 most parsimonious trees of 1602 evolutionary steps with a consistency index (CI) of 0.5306 and a retention index (RI) of 0.7705. As for *trnL-trnF*, 196400 trees of 265 steps were obtained, with a CI of 0.7321 and an RI of 0.8476. With the combined data set, 105 most parsimonious trees of 1855 steps were found (CI of 0.5558 and RI of 0.7672), and their consensus is shown in Fig. 4. The g1 statistic for the ITS trees is −0.992 and for the *trnL-trnF* trees is −1.319, indicating that for both regions the data are significantly skewed and therefore have a substantial phylogenetic signal (Hillis, 1991). For the combined trees, the g1 statistic is −1.058. The maximum-parsimony trees estimated using these parameters were very similar to the respective ML trees (see Supplementary Data accompanying the online version of this article).

The trees obtained by the NJ method with the Kimura-2 parameters and *p* distance were also very similar to the trees obtained with the other methods, especially when the major clades are considered (see Supplementary Data). The NJ trees obtained with the logDet distance did not show significant differences (as assessed by the KH and SH tests) from the others presented here.

There are some differences among the trees on the positioning of a few *Passiflora* species within the major clades that are worth noting. *Passiflora lancetillensis* and *P. microstipula*, from subgenus *Deidamioides*, form a sister group to the *Decaloba* clade in the *trnL-trnF* tree (Fig. 2) with a high divergence and therefore could possibly be considered a fourth main clade. But in the ITS tree (Fig. 1), *P. lancetillensis* is placed within the *Decaloba* clade. *Passiflora rubra* is also placed as a highly divergent sister to the remaining species of

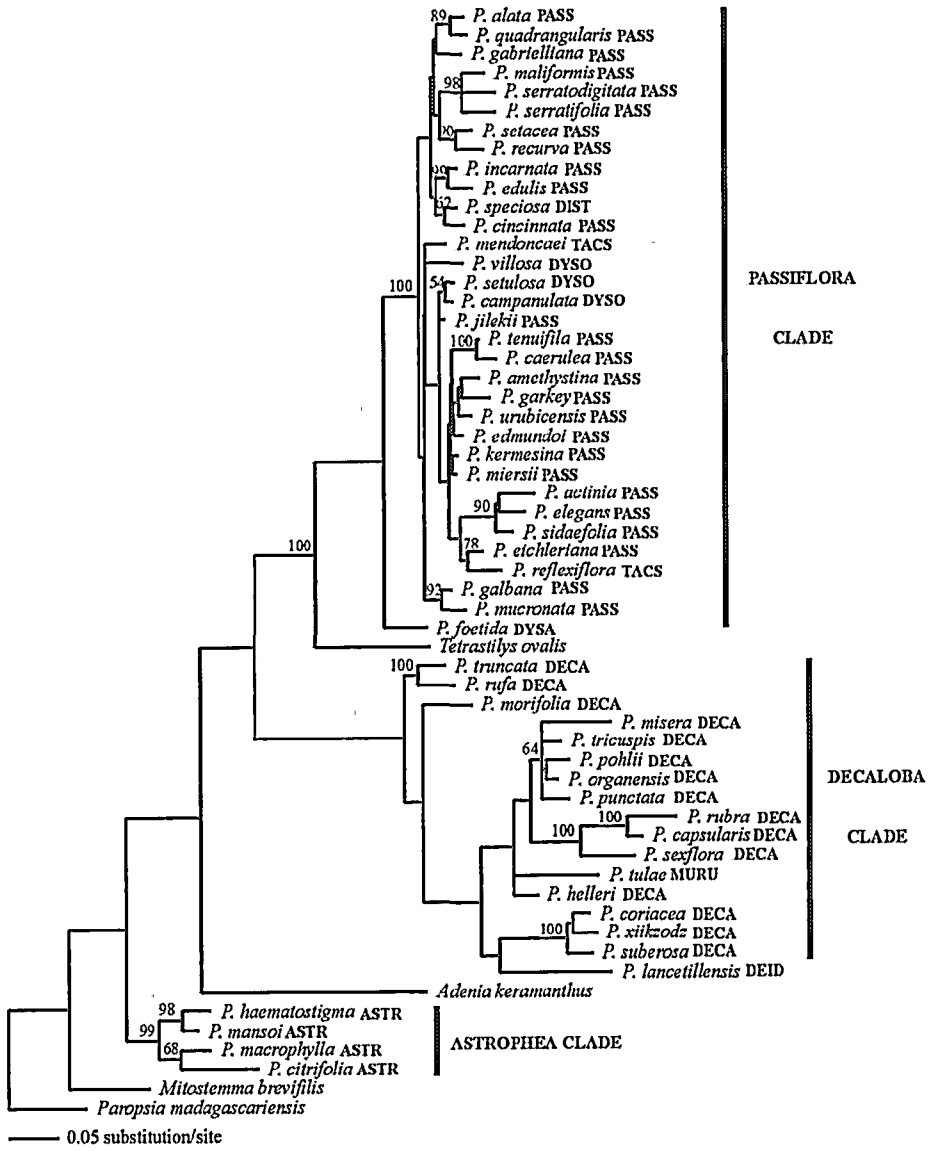

Fig. 3. Maximum-likelihood tree for the ITS + *trnL-trnF* data set in *Passiflora* and outgroups. Numbers above branches are bootstrap support values (when higher than 50%) based on 1000 replicates. Abbreviations indicate the subgenera (see Fig. 1 for full names).

the *Decaloba* clade in the *trnL-trnF* trees (Fig. 2) but not in the ITS trees. Also noteworthy is the position of *P. foetida* in the ITS tree (Fig. 1), as sister to all other taxa in the *Passiflora* clade, a placement not seen in the *trnL-trnF* trees (Fig. 2).

By contrast with the high support among markers and phylogenetic methods for the major *Passiflora* clades, the relationships among the outgroups and clades received very low support (bootstrap and posterior probabilities values) in each tree and with many disagreements among them. For example, in the ITS maximum-likelihood tree (Fig. 1), the outgroups are not phylogenetically adjacent, because *Tetrastilys ovalis* and *Adenia keramanthus* are positioned as sister to the *Passiflora* clade (with a support value of 97%), while *Mitostemma brevifilis* is placed as a sister to the *Decaloba* + *Astrophea* clades (with support of <50%). In the *trnL-trnF* tree (Fig. 2), *T. ovalis* was included within the *Passiflora* clade as sister to *P. speciosa*, but here the other three outgroup species did group together, although with medium (73%) bootstrap support. The combined ML tree (Fig. 3) presents a topology that has components found in both single marker trees, with *T. ovalis* positioned as sister of the *Passiflora* clade (bootstrap support = 100%), *Adenia keramanthus* positioned among the three clades, and *Mitostemma brevifilis* appearing as the most basal species, but both with support values below 50%.

All trees of the *rps4* data set, despite its much smaller size, corroborate the existence of three major *Passiflora* clades, but there is disagreement on the relationship among the clades and their relationship with the outgroups when comparing the different methods. For example, in the ML tree the higher-order relationships are statistically unresolved (Fig. 5), but in the NJ and Bayesian analyses (see Supplementary Data) there is moderate support for the non-monophyly of *Passiflora*. Therefore, the monophyly of the genus *Passiflora* was not statistically supported by any data set or phylogenetic method, and in several trees there was even statistical support for a non-monophyletic *Passiflora*.

The ITS high sequence variability and pairwise distances (see earlier) suggest substitution saturation at higher levels of differentiation. Indeed, this saturation in the ITS region is evident in Fig. 6, in which we plotted the ITS pairwise sequence

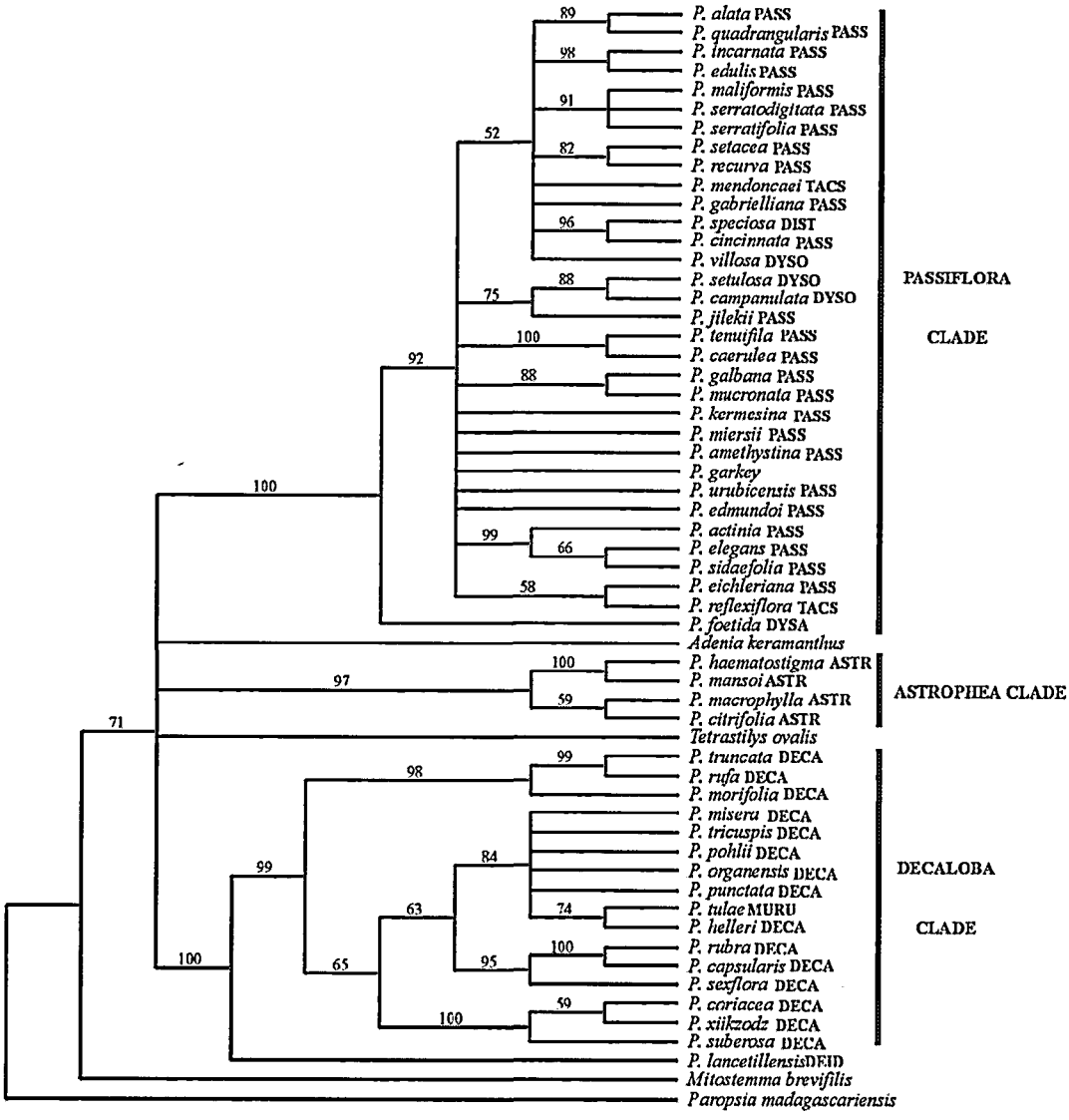

Fig. 4. Strict consensus of the 105 most parsimonious trees for the ITS + *trnL-trnF* data set in *Passiflora* and outgroups. Numbers above branches are bootstrap support values (when higher than 50%) based on 1000 replicates. Abbreviations indicate the subgenera (see Fig. 1 for full names).

p distances with the much more conserved *trnL-trnF* pairwise distances. These graphics suggests that the ITS distances become saturated above a *p* distance of about 0.4, and this saturation implies a low reliability for the highest topological levels of the ITS phylogenetic trees.

The degree of compatibility between the trees obtained with ITS and *trnL-trnF* was different depending on which methods were considered. Despite the general similarity of the trees, mainly at the lower taxonomical levels, Templeton's test and the IDL test detected significant differences among the spacer-region data sets in the parsimony trees. This pattern was similar to those found by Yoder et al. (2001) and Reeves et al. (2001) in other organisms. At present, it is not clear whether heterogeneous sets of data should or should not be combined to obtain an overall phylogenetic picture. On the other hand, for ML and NJ trees, the KH and SH tests did not show significant differences between the spacers.

One interesting difference that can be observed visually is the higher divergence among the species of the *Decaloba* clade (implied by the usually longer branch lengths) as compared to those of the *Passiflora* clade, especially marked for the *trnL-trnF* region (Fig. 2). Also noteworthy is the usually longer branch leading to the *Decaloba* clade, as contrasted to that leading to the *Passiflora* clade. Nucleotide diversity (π) for the ITS region is significantly higher in the *Decaloba* clade (0.102 ± 0.013) as compared to the *Passiflora* clade (0.043 ± 0.078). The relative rate *Z* statistic ($Z = 3.549$) calculated with the PHYLTEST package rejects rate constancy ($P < 0.05$) between the two. A similar result was found for the *trnL-trnF* intergenic region ($Z = 2.399$, $P < 0.05$) and for the *rps4* gene ($Z = 2.976$, $P < 0.05$), suggesting that these different evolutionary rates are not locus specific but characteristic of the clades. To test if these differences were not due to the different number of species sampled in them, new analyses were performed consisting of 10 random subsamples of species with an identical number of species in both clades. The results corroborated the significant differences between the two clades. Because of its much smaller sample size, we did not test the *Astrophea* clade for these divergence differences.

Anyone acquainted with *Passiflora* could intuitively perceive that the different subgenera could also have dissimilar flower sizes. Indeed, a Wilcoxon-Mann-Whitney test showed

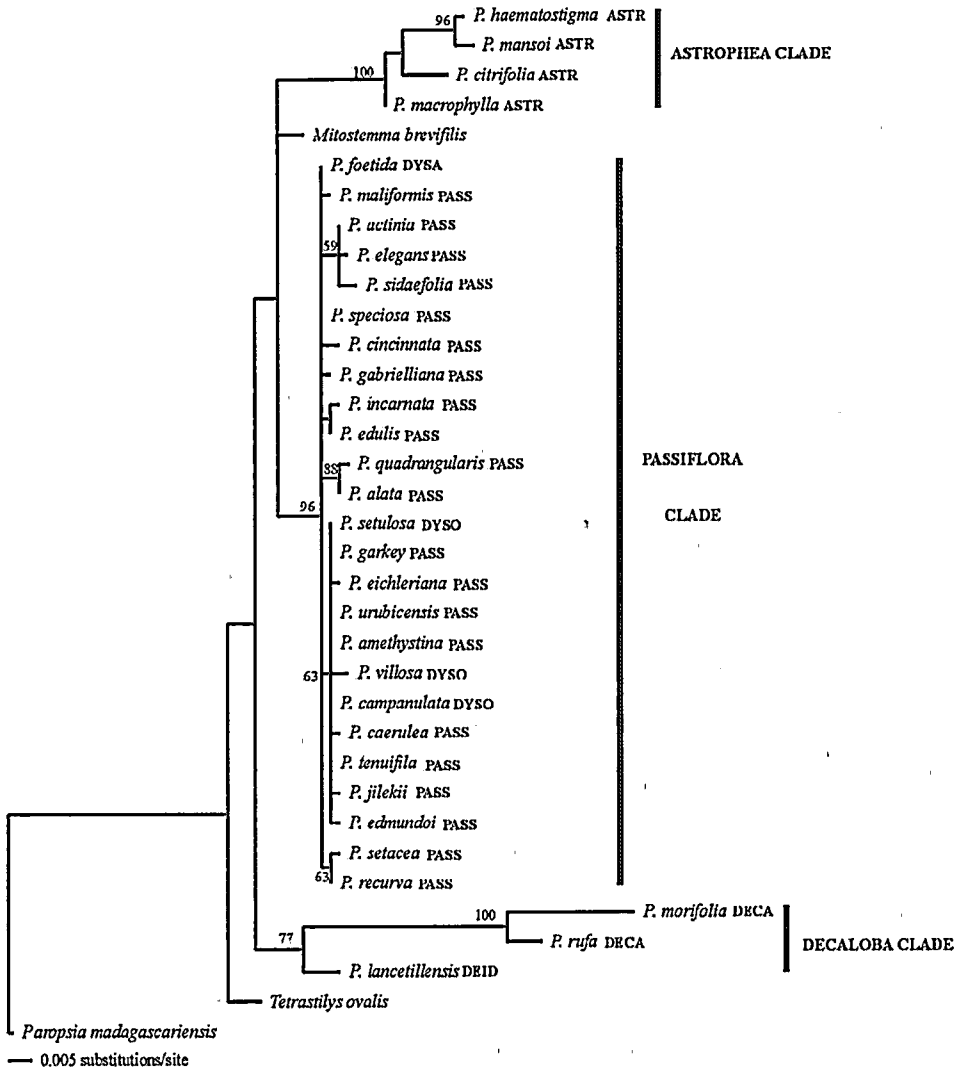

Fig. 5. Maximum-likelihood tree for the *rps4* gene in *Passiflora* and outgroups. Numbers above branches are bootstrap support values (when higher than 50%) based on 1000 replicates. Abbreviations indicate the subgenera (see Fig. 1 for full names).

that the flowers of the *Passiflora* clade have a significantly larger ($P < 0.001$) minimum size than the flowers of the *Decaloba* and *Astrophea* clades (median values: 19 mm vs. 6.5 mm, respectively).

DISCUSSION

Taxonomic implications—No data set or method generates a phylogenetic tree that is consistent with *Passiflora*'s monophyly. However, there is much inter-tree variability in the relationships between the outgroups and the *Passiflora* clades, and most of the groupings received low branch support. Considering these results, additional data would be necessary to test the monophyly of the genus *Passiflora* and its relationships with the other genera of the family.

By contrast, all phylogenetic results presented here suggest the existence of three well-defined clades in *Passiflora*. The largest comprise all representatives of Killip's (1938) subgenera *Distephana*, *Dysosmia*, *Dysosmioides*, *Passiflora*, and *Tacsonioides*. Because the subgenus *Passiflora* is by far the most diverse of all, we named this set as the *Passiflora* clade. A second major clade comprises all species of Killip's (1938) and MacDougal's (1994) subgenus *Decaloba*, as well *P. tulae* of subgenus *Murucuja*, *P. cuprea* of subgenus *Pseudomurucuja*, and *P. multiflora* of subgenus *Adopogyne*. Using the same reasoning stated above, we called this the *Decaloba* clade. The third clade proposed here comprises the four species of the subgenus *Astrophea* (*P. haematostigma*, *P. mansoi*, *P. citrifolia*, and *P. macrophylla*), which has very high support values in most of the trees presented here.

The subgenus *Deidamioides* (represented here by *P. lancetillensis* and *P. microstipula*) presents ambiguous positions in the trees derived from the different markers. In the more conserved *trnL-trnF* spacer and *rps4* gene, this subgenus is very divergent from the other subgenera, and it was placed in some *trnL-trnF* trees (e.g., the MP tree available in the supplementary data) as a very basal independent clade, suggesting that it could be considered as a fourth major clade of *Passiflora*. However, in the ITS trees (e.g., Fig. 1), *P. lancetillensis* is placed within the *Decaloba* clade, with a high bootstrap value. These ambiguities, and the low number of species sampled, determine that the status of this subgenus should remain undefined until further data are analyzed.

Though we could not find any published evidence that explicitly corroborates the existence of these three clades in *Passiflora*, two papers seem to agree with our results. De Melo et

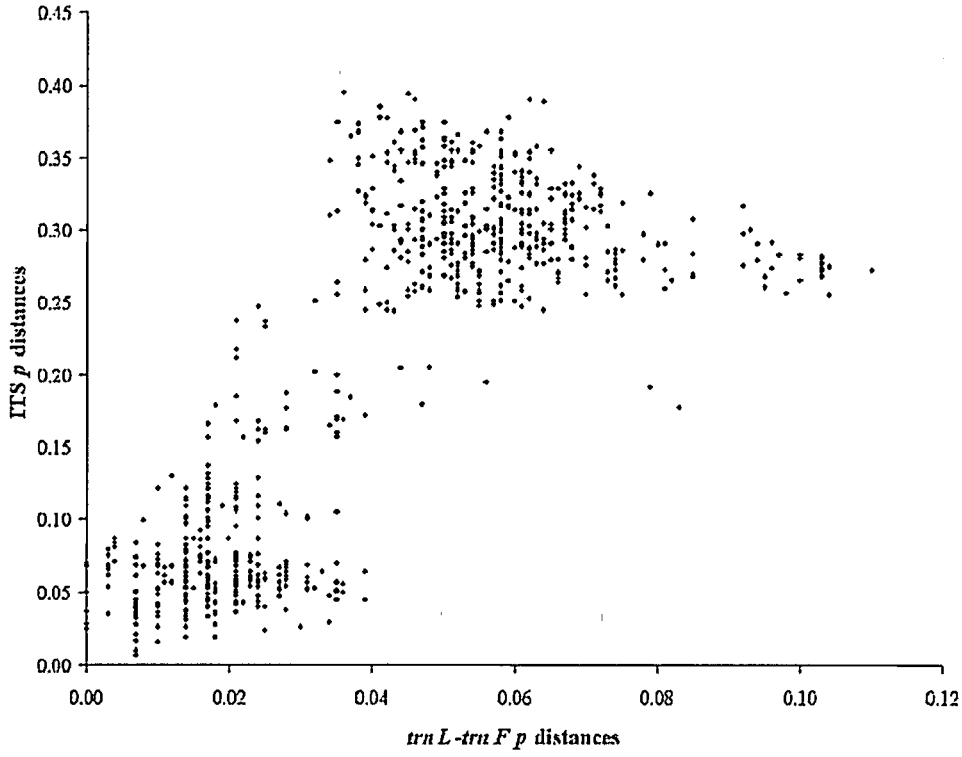

Fig. 6. Scatterplot of the relationship between ITS and *trnL-trnF* pairwise sequence distances (*p* distances) among all pairs of sequences in *Passiflora* and outgroups. The ITS distances became saturated above a *p* distance of about 0.4.

al. (2001) described the basic chromosome number of several subgenera of *Passiflora* and some related genera. From the interpretation of their tree (their Fig. 6) in a phylogenetic context, a grouping of subgenera was suggested that is consistent with the three clades described here. On the other hand, Downie et al. (1996) studied the *rpoC1* chloroplast intron in several families and found that it was absent in four species of the subgenus *Decaloba*, but present in six species of the *Passiflora* clade (three subgenera). It should also be remembered (see Results) that the five subgenera that form the *Passiflora* clade all have statistically significant larger and more conspicuous flowers than those of the *Decaloba* and *Astrophea* clades.

Although there are some consistent species groupings within the *Passiflora* and *Decaloba* clades, only a few of them have high levels of statistical support and are consistent among markers and phylogenetic methods, such as the *P. actinia* + *P. elegans* + *P. sidaefolia* and the *P. quadrangularis* + *P. alata* subgroups. Most of these subgroups, however, are not consistent with Killip's (1938) subgenera or sections.

One question that can be considered is the taxon sampling of our study and its bearing on the reliability of the major clades found. We are aware that a larger sample of species could result in more reliable trees. However, we suggest that our results and phylogenetic conclusions are presently adequate. Thus, our sampling represents about 15% and 49% of *Passiflora*'s species and subgenera, respectively. These sample sizes are by no means unusual or small in phylogenetic studies of large and diversified taxa. For example, Miller et al. (2002) investigated the monophyly of *Ipomoea* (Convolvulaceae) using just 35 (5.8%) of about 600 species of this genus; Bortiri et al. (2002) studied just 22 (11%) of over 200 species of *Prunus*; and Cuénoud et al. (2002) studied just one species of each genus for the phylogenetic analyses of Caryophyllales. Perhaps the clearest example of the utility of a limited taxonomic sampling is the most recent work that tries to infer ordinal and familial classification for angiosperms (such as that of the Angiosperm Phylogeny Group, 1998) with a relatively limited sample of the angiosperm species. Rosenberg and Kumar (2001) also suggested that incomplete sampling might not be a problem for phylogenetic inferences (but see Zwickl and Hillis [2002] for a different position). Our results with 11 (49%) of the 23 putative subgenera suggest that at least the subgenera studied here should be grouped in three major clades: *Passiflora*, *Decaloba*, and *Astrophea*.

Sequence variation—The GC content in the ITS region varied from 47% to 67% in the different species of *Passiflora*, a variability that has also been observed in other groups of plants (e.g., Nishikawa et al., 1999; Gaut et al., 2000). Interestingly, the GC content differs significantly between the *Decaloba* and non-*Decaloba* clades, and although this may be a bias factor for most phylogenetic methods, the use of the logDet distance (Lockhart et al., 1994; Steel, 1994) should correct for it. The use of this distance with a subsample of our data yielded the same main clades presented here, suggesting that the ITS phylogeny was not biased by the heterogeneity in nucleotide content (see Supplementary Data).

Heterogeneity in the rates of DNA change such as that observed in the *Decaloba*/non-*Decaloba* clades has also been documented at several taxonomic levels in plants (Bousquet et al., 1992; Gaut et al., 1993; Suh et al., 1993; Gielly and Taberlet, 1996; Eyre-Walker and Gaut, 1997; Aïnouche and Bayer, 1999; Bortiri et al., 2002). Many factors supposedly influence the rate of molecular evolution, such as generation time, selection, DNA replication or repair, and metabolic rates (Wu and Li, 1985; Britten, 1986; Bousquet et al., 1992; Li, 1993; Martin and Palumbi, 1993; Eyre-Walker and Gaut, 1997; Gaut et al., 1997). Unfortunately, little is known about

such factors in *Passiflora*, although Benson at al. (1975) asserted that members of the subgenus *Decaloba* have shorter generation times than species of the subgenus *Passiflora*, a factor that may have accelerated its evolutionary rate. Another possible explanation for the very short branches within the *Passiflora* clade is a rapid radiation of the species of this clade, an explanatory factor that was attributed to similar results described in other groups of angiosperms (e.g., Bortiri et al., 2002).

Comparison among markers—The discrepancies between the trees obtained with ITS and *trnL-trnF* are mainly due to the position of a few species (e.g., *P. rubra* and *P. foetida*) and the outgroup relationships. Some of these discrepancies could be due to the different forms of inheritance of nuclear and chloroplast markers. The majority of the angiosperms have maternal chloroplast inheritance, but *Turnera ulmifolia*, a closely related member of the Malpighiales, has mixed, paternally biased inheritance (Shore and Triassi, 1998). (Note that Chase et al. [2002] placed the Turneraceae within the Passifloraceae.) Courriveau and Coleman (1988) found chloroplast paternal inheritance in *P. edulis*; and similar results were obtained by A. P. Lorenz et al. (Federal University of Rio Grande do Sul, unpublished data) in *P. elegans* and *P. actinia*. In accordance with this reasoning, the trees based on the two chloroplast markers, *trnL-trnF* and *rps4*, generally had fewer discrepancies among themselves than when they were compared with trees based on ITS sequences.

Another factor that may have contributed to these discrepancies, mainly at the higher level relationships, is the observed saturation in the ITS region, when contrasted to the *trnL-trnF* data (Fig. 5). The phylogenetic utility of the ITS spacers within the Passifloraceae seems to be restricted to the interspecific comparisons, and they should be used with caution at the intergeneric level.

Future considerations—We recommend, and it is in our plans, that future work should concentrate on certain critical areas, such as increased taxon sampling, mainly of the subgenera not yet studied and of the larger subgenera, such as *Decaloba*, that were underrepresented here. Furthermore, considering the lack of adequate resolution found in our trees at the higher order relationships, it is necessary to increase the amount of sequence data to clarify issues such as the monophyly of *Passiflora* and its relationship with its major clades and with the other genera of the Passifloraceae. Also of interest is an understanding of the causes underlying the different evolutionary dynamics found here among the major *Passiflora* clades.

LITERATURE CITED

AÏNOUCHE, A.-K., AND R. J. BAYER. 1999. Phylogenetic relationships in *Lupinus* (Fabaceae: Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. *American Journal of Botany* 86: 590–607.

ANGIOSPERM PHYLOGENY GROUP. 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85: 531–553.

BAKKER, F. T., D. HELLBRÜGGE, A. CULHAM, AND M. GIBBY. 1998. Phylogenetic relationships within Pelargonium sect. Peristera (Geraniaceae) inferred from nrDNA and cpDNA sequence comparisons. *Plant Systematics and Evolution* 211: 273–287.

BALDWIN, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3–16.

BALDWIN, B. G., M. J. SANDERSON, M. J. 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.

BENSON, W. W., K. S. BROWN, AND L. E. GILBERT. 1975. Coevolution of plants and herbivores: passion flower butterflies. *Evolution* 29: 659–680.

BORTIRI, E., S.-H. OH, F.-Y. GAO, AND D. POTTER. 2002. The phylogenetic utility of nucleotide sequences of sorbitol 6-phosphate dehydrogenase in *Prunus* (Rosaceae). *American Journal of Botany* 89: 1697–1708.

BOUSQUET, J. S., A. H. STRAUSS, A. H. DOERKSEN, AND R. A. PRICE. 1992. Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. *Proceedings of National Academy of Sciences USA* 89: 7844–7848.

BRITTEN, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231: 1393–1398.

CERVI, A. C. 1997. Passifloraceae do Brasil. Estudo do gênero *Passiflora* L., subgênero *Passiflora*. *Fontqueria* 45: 1–92.

CHASE, M. W., S. ZMARZTY, M. D. LLEDÓ, K. J. WURDACK, S. M. SWENSEN, AND M. F. FAY. 2002. When in doubt, put it in Flacourtiaceae: a molecular phylogenetic analysis based on plastid *rbcL* DNA sequences. *Kew Bulletin* 57: 141–181.

COURRIVEAU, J. L., AND A. W. COLEMAN. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *American Journal of Botany* 75: 1443–1458.

CUÉNOUD, P., V. SAVOLAINEN, L. W., M. POWELL, R. J. GRAYER, AND M. W. CHASE. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbc*L, *atp*B, and *mat*K DNA sequences. *American Journal of Botany* 89: 132–144.

DE MELO, N. F., A. C. CERVI, AND M. GUERRA. 2001. Karyology and cytotaxonomy of the genus *Passiflora* L. (Passifloraceae). *Plant Systematics and Evolution* 226: 69–84.

DESFEUX, C., AND B. LEJEUNE. 1996. Systematics of Euromediterranean *Silene* (Caryophyllaceae): evidence from a phylogenetic analysis using ITS sequence. *Comptes Rendus, Academie des Sciences, Paris* 319: 351–358.

DOWNIE, S. R., E. LLANAS, AND D. S. KATZ-DOWNIE. 1996. Multiple independent losses of the *rpoC1* intron in angiosperm chloroplast DNA's. *Systematic Botany* 21: 135–151.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA leaf material. *Phytochemical Bulletin* 19: 11–15.

ESCOBAR, L. K. 1988. Passifloraceae. *In* P. Pinto and G. Lozano [eds.], Flora de Colombia, vol. 10, 1–138. Universidad Nacional de Colombia, Bogotá, Colombia.

ESCOBAR, L. K. 1989. A new subgenus and five new species in *Passiflora* (Passifloraceae) from South America. *Annals of the Missouri Botanical Garden* 76: 877–885.

EYRE-WALKER, A., AND B. S. GAUT. 1997. Correlated rates of synonymous site evolution across plant genomes. *Molecular Biology and Evolution* 14: 455–460.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1994. Testing significance of congruence. *Cladistics* 10: 315–319.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

GAUT, B. S., L. G. CLARK, J. F. WENDEL, AND S. V. MUSE. 1997. Comparisons of the molecular evolutionary process at *rbc*L and *ndh*F in the grass family (Poaceae). *Molecular Biology and Evolution* 14: 769–777.

GAUT, B. S., S. V. MUSE, AND M. T. CLEGG. 1993. Relative rate of nucleotide substitution in the chloroplast genome. *Molecular Phylogenetics and Evolution* 2: 89–96.

GAUT, B. S., L. P. TREDWAY, C. KUBIK, R. L. GAUT, AND W. MEYER. 2000. Phylogenetic relationships and genetic diversity among members of the *Festuca-Lolium* complex (Poaceae) based on ITS sequence data. *Plant Systematics and Evolution* 224: 33–53.

GIELLY, L., AND P. TABERLET. 1996. A phylogeny of the European gentians inferred from chloroplast *trnL* (UAA) intron sequences. *Botanical Journal of Linnean Society* 120: 57–75.

HARDIG, T. M., P. S. SOLTIS, AND D. E. SOLTIS. 2000. Diversification of the North American shrub genus *Ceanothus* (Rhamnaceae): conflicting phylogenies from nuclear risbosomal DNA and chloroplast DNA. *American Journal of Botany* 87: 108–123.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human-ape

splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160–174.

HEDGES, S. B. 1992. The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies. *Molecular Biology and Evolution* 9: 366–369.

HILLIS, D. M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. *In* M. M. Myamoto and J. Craft [eds.], Phylogenetic analysis of DNA sequences, 278–294. Oxford University Press, New York, New York, USA.

HUELSENBECK, J. P., AND F. RONQUIST. 2001. MrBayes: Bayesian inference of phylogeny. Department of Biology, University of Rochester, Rochester, New York, USA.

JUDD, W. S., C. S. CAMPBELL, E. A. KELLOGG, AND P. F. STEVENS. 1999. Plant systematics: a phylogenetic approach. Sinauer Associates, Sunderland, Massachusetts, USA.

KILLIP, E. P. 1938. The American species of Passifloraceae. *Field Museum of Natural History, Botanical Series* 19: 1–613.

KISHINO, H., AND M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topology from DNA sequence data, and the branching order in Hominoidea. *Journal of Molecular Evolution* 29: 170–179.

KUMAR, S. 1996. PHYLTEST: phylogeny hypothesis testing software. Pennsylvania State University, University Park, Pennsylvania, USA.

LEMMON, A., AND M. C. MILINKOVITCH. 2002. The metapopulation genetic algorithm: an efficient solution for problem of large phylogeny estimation. *Proceedings of National Academy of Sciences, USA* 99: 10516–10521.

LI, W. H. 1993. So, what about the molecular clock hypothesis? *Current Opinion in Genetics and Development* 3: 896–901.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, AND D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution* 11: 605–612.

MACDOUGAL, J. M. 1994. Revision of *Passiflora* subgenus *Decaloba* section Pseudodysosmia (Passifloraceae). *Systematic Botany Monographs* 41: 1–146.

MARTIN, A. P., AND S. R. PALUMBI. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of National Academy of Sciences, USA* 90: 4087–4091.

MILLER, R. E., T. R. BUCKLEY, AND P. S. MANOS. 2002. An examination of monophyly of morning glory taxa using Bayesian phylogenetic inference. *Systematic Biology* 51: 740–753.

MOLVRAY, M., P. J. KORES, AND M. W. CHASE. 1999. Phylogenetic relationships within Korthalsella (Viscaceae) based on nuclear ITS and plastid *trnL-F* sequence data. *American Journal of Botany* 86: 249–260.

NADOT, S., R. BAJON, AND B. LEJEUNE. 1994. The chloroplast gene *rps*4 as a tool for the study of Poaceae phylogeny. *Plant Systematics and Evolution* 191: 27–38.

NEI, M., AND S. KUMAR. 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford, UK.

NISHIKAWA, T., K. OKAZAKI, T. UCHINO, K. ARAKAWA, AND T. NAGAMINE. 1999. A molecular phylogeny of *Lilium* in the internal transcribed spacer region of nuclear ribosomal DNA. *Journal of Molecular Evolution* 49: 238–249.

POSADA, D., AND K. A. CRANDALL. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

RAUSCHER, J. T. 2002. Molecular phylogenetics of the *Espletia* complex (Asteraceae): evidence from nrDNA ITS sequences on the closest relatives of Andean adaptive radiation. *American Journal of Botany* 98: 1074–1084.

REEVES, G., M. W. CHASE, P. GOLDBLATT, P. RUDALL, M. F. FAY, A. V. COX, B. LEJEUNE, AND T. SOUZA-CHIES. 2001. Molecular systematics of Iridaceae: evidence from four plastid regions. *American Journal of Botany* 88: 2074–2087.

ROSENBERG, M. S., AND S. KUMAR. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. *Proceedings of National Academy of Sciences, USA* 98: 10 751–10 756.

SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 9: 945–967.

SÁNCHEZ, I., F. ANGEL, M. GRUMM, M. C. DUQUE, M. LOBO, J. TOHME, AND W. ROCA. 1999. Variability of chloroplast DNA in the genus *Passiflora* L. *Euphytica* 106: 15–26.

SHIMODAIRA, H., AND M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* 16: 1114–1116.

SHORE, J. S., AND M. TRIASSI. 1998. Paternally biased cpDNA inheritance in *Turnera ulnifolia* (Turneraceae). *American Journal of Botany* 85: 328–332.

SOUZA-CHIES, T. T., G. BITTAR, S. NADOT, L. CARTER, E. BESIN, AND B. LEJEUNE. 1997. Phylogenetic analysis of *Iridaceae* with parsimony and distance methods using the plastid gene *rps*4. *Plant Systematics and Evolution* 204: 109–123.

STEEL, M. 1994. Recovering a tree from the Markov leaf colouration it generates under a Markov model. *Applied Mathematics Letters* 7: 19–23.

SUH, Y., L. B. THIEN, H. E. REEVE, AND E. A. ZIMMER. 1993. Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. *American Journal of Botany* 80: 1042–1055.

SUZUZI, Y., G. V. GLAZCO, AND M. NEI. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of National Academy Sciences, USA* 99: 16 138–16 143.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts, USA.

TABERLET, P., L. GIELLY, G. PAUTOU, AND J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

TAKEZAKI, N. 2001. SendBS Program. Available at website: www.cib.nig.ac.jp/dda/ntakezak.html#sendbs.

TAKEZAKI, N., A. RZHETSKY, AND M. NEI. 1995. Phylogenetic test of the molecular clock and linearized tree. *Molecular Phylogenetics and Evolution* 12: 823–833.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 24: 218–227.

THOMPSON, J. D., D. G. HIGGINS, AND T. J. GIBSON. 1994. ClustalW: improving the sensivity of progressive multiple sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

THOMPSON, J. D., D. G. HIGGINS, AND T. J. GIBSON. 2001. ClustalX. Program available at ftp://ftp-igbmc.u-strasbg.fr/pub/clustalx/.

WILCOX, T. P., D. J. ZWICKL, T. A. HEATH, AND D. M. HILLIS. 2002. Phylogenetic relationships of dwarf boas and comparison of Bayesian and bootstrap measures of phylogenetic support. *Molecular Phylogenetics and Evolution* 25: 361–371.

WU, C. I., AND W. H. LI. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proceedings of National Academy of Sciences, USA* 82: 1741–1745.

XIANG, Q.-Y., M. L. MOODY, D. E. SOLTIS, C. FAN, AND P. S. SOLTIS. 2002. Relationships within Cornales and circumscription of Cornaceae—*mat*K and *rbc*L sequence data and effects of outgroups and long branches. *Molecular Phylogenetics and Evolution* 24: 35–57.

YODER, A. D., J. A. IRWIN, AND B. A. PAYSEUR. 2001. Failure of the ILD test to determine data combinability for slow loris phylogeny. *Systematic Biology* 50: 408–424.

ZWICKL, D. J., AND D. M. HILLIS. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology* 51: 588–598.