*Mycologia,* 103(2), 2011, pp. 361–378. DOI: 10.3852/10-120


# Align or not to align? Resolving species complexes within the *Caloplaca saxicola* group as a case study

Ester Gaya[1]

*Department of Plant Biology (Botany Unit), Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, 08028 Barcelona, Spain*

Benjamin D. Redelings

*National Evolutionary Synthesis Center (NESCent), Durham, North Carolina 27705*

Pere Navarro-Rosinés
Xavier Llimona

*Department of Plant Biology (Botany Unit), Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, 08028 Barcelona, Spain*

Miquel De Cáceres

*Biodiversity and Landscape Ecology Lab, Centre Tecnològic Forestal de Catalunya, Ctra. St. Llorenç de Morunys km 2, 25280 Solsona, Spain*

François Lutzoni

*Department of Biology, Duke University, Durham, North Carolina 27708-0338*

***Abstract:*** The *Caloplaca saxicola* group is the main group of saxicolous, lobed-effigurate species within genus *Caloplaca* (Teloschistaceae, lichen-forming Ascomycota). A recent monographic revision by the first author detected a wide range of morphological variation. To confront the phenotypically based circumscription of these taxa and to resolve their relationships morphological and ITS rDNA data were obtained for 56 individuals representing eight *Caloplaca* species belonging to the *C. saxicola* group. We tested the monophyly of these eight morphospecies by performing maximum parsimony, maximum likelihood and two different types of Bayesian analyses (with and without a priori alignments). Restricting phylogenetic analyses to unambiguously aligned portions of ITS was sufficient to resolve, with high bootstrap support, five of the eight previously recognized species within the *C. saxicola* group. However, phylogenetic resolution of all or most of the eight species currently included as two distinct subgroups within the *C. saxicola* group was possible only by combining morphological characters and signal from ambiguously aligned regions with the unambiguously aligned ITS sites or when the entire ITS1 and 2 regions were not aligned a priori and included as an integral component of a Bayesian analysis (BAli-Phy). The *C. arnoldii* subgroup includes *C. arnoldii*, comprising four subspecies, and the *C. saxicola* subgroup encompasses seven species. Contrary to the *C. saxicola* subgroup, monophyly of taxa included within the *C. arnoldii* subgroup and their relationships could not be resolved with combined ITS and morphological data. Unequivocal morphological synapomorphies for all species except *C. arnoldii* and *C. pusilla* are recognized and presented.

***Key words:*** *Caloplaca*, *Gasparrinia*, lichens, molecular phylogeny, morphological phylogeny, morphospecies

## INTRODUCTION

The widespread lichen-forming genus *Caloplaca* Th. Fr. is the largest of the Teloschistaceae, with about 500 known species (Kirk et al. 2008). It includes mostly crustose species that often have anthraquinones in both thalli and apothecia and colorless polarilocular ascospores. The delimitation of *Caloplaca* and its relationship to other genera of Teloschistaceae is controversial and has been the objective of several phylogenetic studies (Arup and Grube 1999; Gaya et al. 2003, 2008; Søchting and Lutzoni 2003). Furthermore, problems of delimitation are not restricted to genus but also occur among species groups and taxa within *Caloplaca*. Specifically, the lobed *Caloplaca* species constitute a difficult group encompassing extraordinary phenotypic variation. Clauzade and Roux (1985) referred to these lobed *Caloplaca* as subgenus *Gasparrinia* (Torn.) Th. Fr. and defined within it five groups: *C. aurantia*, *C. aurea*, *C. carphinea*, *C. persica* and *C. saxicola*. The *C. saxicola* group, which represents the core of the saxicolous lobed-effigurate *Caloplaca* species, is among the most taxonomically controversial. However not all authors have followed Clauzade's and Roux's classification and often considered the lobed *Caloplaca* species as a whole, without segregating them into groups (e.g. Verseghy 1970, 1971, 1972; Nordin 1972; Kärnefelt 1989; Wetmore and Kärnefelt 1998; see also Gaya et al. 2008).

Adding to the taxonomic complexity, some of the taxa treated here often have been included within a broader concept of a single taxon, *C. saxicola*

Submitted 21 Apr 2010; accepted for publication 27 Aug 2010.

[1] Corresponding author. Current address: Department of Biology, Duke University, Durham, North Carolina 27708-0338. E-mail: eb62@duke.edu

(Hoffm.) Nordin. Thus Poelt (1954) distinguished what he called *C. arnoldii* s.l., including *C. arnoldii-confusa* Gaya & Nav.-Ros., from *C. murorum* (Hoffm.) Th. Fr. (= *C. saxicola*), but included within *C. murorum* what is currently recognized as *C. arnoldii* ssp. *obliterata* (Pers.) Gaya and *C. rouxii* Gaya, Nav.-Ros. & Llimona. However, he took into account the variability of *C. murorum* and informally used several names to refer to the diversity of morphologies associated with different ecologies and suggested the possibility of detecting truly genetically based forms (independent from environmental factors) in future studies (Gaya 2009). Nordin (1972) later synonymized *C. murorum* f. *miniata* (Hoffm.) Ozenda & Clauzade (currently *C. rouxii*), some of the morphotypes now included within *C. arnoldii* ssp. *obliterata* (e.g. ''*obliterata*,'' ''*obliterascens*'') and *C. pusilla* (A. Massal.) Zahlbr. under *C. saxicola*. Apart from *C. decipiens* (Arnold) Blomb. & Forssell, Wetmore and Kärnefelt (1998) only recognized *C. saxicola* and described three morphotypes within it. The conflicts between these taxonomic treatments are understandable because, according to Gaya (2009), a wide range of morphological variation at both the intra- and interspecific levels is present in this group, causing major problems for species delimitation and identification, especially for taxa within the *C. arnoldii* complex (part of the *C. saxicola* group), where intermediate phenotypes are frequently observed within a population or even within a single thallus.

Phylogenetic studies on the Teloschistaceae (Arup and Grube 1999; Gaya et al. 2003, 2008; Søchting and Lutzoni 2003) have confirmed that genus *Caloplaca* and subgenus *Gasparrinia* are both polyphyletic. Furthermore, the *C. saxicola* group sensu Clauzade and Roux (1985) includes monophyletic groups resulting from several independent origins (Gaya et al. 2003, 2008). Gaya (2009) restricted the *C. saxicola* group to the largest clade, which includes *C. arnoldii*, *C. arnoldiiconfusa*, *C. biatorina* (A. Massal.) J. Steiner, *C. decipiens*, *C. pusilla*, *C. rouxii*, *C. saxicola* s. str. and *C. schistidii* (Anzi) Zahlbr., and defines the focal taxonomic group of this study. All previous taxonomic studies of the *C. saxicola* group were based exclusively on phenotypic characters. In general, only a few molecular investigations have focused on species complexes within *Caloplaca*; Arup (2006) and Vondrák et al. (2009) focused on the *C. citrina* complex, Arup (2009) on the *C. holocarpa* complex, Muggia et al. (2008) on endolithic taxa of section *Pyrenodesmia* and Vondrák et al. (2008) on the *C. cerina* group. All these studies used the ITS region as a molecular marker.

The main goal of this study was to elaborate an integrated phylogenetic analysis of morphological and molecular data for members of the *C. saxicola* group and to assess the value of morphological characters in delimiting species within species complexes. We conducted a detailed reassessment of morphological and anatomical features and sequenced the nuclear ribosomal internal transcribed spacer (ITS) region of 56 specimens representing 11 taxa within the *C. saxicola* group, which were the subject of a taxonomic revision by the first author (Gaya 2009). Because a high rate of indels is integral to the evolution of the ITS, resulting in the fastest evolving sites being not alignable, much of the signal from ITS1 and 2 must be excluded from conventional phylogenetic analyses, even when the phylogenetic signal is not saturated. Yet to resolve species complexes it is most often these fast-evolving sites that are the most phylogenetically informative. It is possible to capture this signal by adopting a coding procedure that does not require multiple sequence alignments or by using methods in which the alignment is one of the variables optimized during the phylogenetic search. Hence a second objective of this study was to compare the abilities of two methods to resolve species complexes with ITS data. A novel Bayesian method that does not require sequences to be aligned or coded was compared with a second approach in which the signal from nonalignable regions is coded and added to the matrix of alignable sites before conducting an unequally maximum parsimony analysis.

## MATERIALS AND METHODS

*Taxon sampling.*—This study was carried out on collections from the first author and on material from selected herbaria (BCN, E, GZU, LD, MARSSJ, MIN, VAB-LICH and the U. Arup personal herbarium). For the phylogenetic analysis based on morphological data we selected a total of 56 specimens that were chosen to represent as much of the range of morphological variation of each taxon as possible. We attempted to keep the number of specimens per taxon proportional to the morphological variation of each taxon, and when possible we included fertile specimens. Nevertheless, due to the scarcity of available material, some taxa are represented only by single specimens. Thus we included six specimens of *Caloplaca arnoldii* ssp. *arnoldii*, one specimen of *C. arnoldii* ssp. *clauzadeana* Gaya, one specimen of *C. arnoldii* ssp. *nana* Gaya, 11 specimens of *C. arnoldii* ssp. *obliterata*, nine specimens of *C. biatorina*, two specimens of *C. decipiens*, six specimens of *C. arnoldiiconfusa*, six specimens of *C. pusilla*, three specimens of *C. rouxii*, 10 specimens of *C. saxicola* s. str. and one specimen of *C. schistidii*. To root the ingroup we included four additional species: *C. ignea* Arup; *Xanthoria candelaria* (L.) Th. Fr.; *X. elegans* (Link.) Th. Fr.; and *X. sorediata* (Vain.) Poelt, for a total of 60 specimens from 12 species (SUPPLEMENTARY TABLE I). The morphological dataset was assembled based on these 60 specimens and subjected to

phylogenetic analyses, following the exemplar approach of Prendini (2001).

For the molecular phylogenetic analyses we used a total of 62 specimens of the *C. saxicola* group (ingroup), including all specimens used in the morphological study. For each taxon we tried to sequence a minimum of two specimens representing geographically distant populations. We selected the outgroup by comparing our alignment of the *C. saxicola* group with sequences of other species of *Caloplaca* and *Xanthoria* and by considering the phylogenetic relationships established in more inclusive trees (Gaya et al. 2003, 2008). Based on these comparisons, we selected six outgroup specimens representing four taxa: *Caloplaca ignea* (two specimens, one from GenBank, Arup and Grube 1999, and one from Gaya et al. 2008); *Xanthoria elegans* (one specimen, Gaya et al. 2008); *X. candelaria* (one specimen, Gaya et al. 2008); *X. sorediata* (two specimens, one from GenBank, Lohtander et al. 2000, and one from Gaya et al. 2008). Two data matrices were prepared: one including the 68 sequences and another with the selected 60 sequences to be used in the combined analysis with the morphological data.

*Character sampling.*—A total of 57 morphological characters were evaluated, of which 19 characters were excluded for at least one of these three reasons: character invariance among taxa; character unreliability due to examination difficulties; excess of character variation in a single specimen. Hence for the phylogenetic analyses we retained 38 characters (SUPPLEMENTARY MATERIALS AND METHODS, SUPPLEMENTARY TABLE II), of which most describe aspects of thalli and apothecia. A single ecological character (type of substrate) also was included in the data matrix, even though it is not strictly morphological, because we consider it taxonomically and evolutionarily important. Of these 38 characters 26 were discrete and 12 continuous. Discrete characters were included directly in the data matrix with MacClade 4.01 (Maddison and Maddison 2001).

Continuous characters first were converted to a discrete form via the *K-means* method (MacQueen 1967), which is a non-hierarchical, polythetic and partitive clustering algorithm. This technique of cluster analysis partitions *K* groups, or clusters, within a set of objects for which we have multivariate information. Its algorithm is based on the iteration of these steps: (i) the positions to the averages (geometric centers) of the *K* clusters are calculated; (ii) for each object, its distance to the *K* averages is calculated; and (iii) each object is reassigned to the nearest cluster. The final clusters are expected to have minimum dispersion (sum of squared deviations from the cluster average), but because the method is iterative several runs are normally performed starting from random configurations. The number of clusters sought ($K$) is a parameter of the method that has to be set a priori. One of the most frequently used statistics designed to help determine the optimum number of groups according to the data at hand is the *Pseudo-F* statistic (Calinski and Harabasz 1974). *Pseudo-F* compares the dispersion among groups with respect to the dispersion within groups, in a similar way to an ANOVA $F$ statistic. The more internally homogeneous the groups, and more differentiated among them, the larger the value of the statistic.

Based on the mean matrices of each character, we applied the *K-means* method on each of the continuous characters separately. For each character we ran the clustering algorithm trying a range of values for $K$, from $K = 2$ to $K = 10$. The resulting profiles were graphically explored to select the number of clusters that show a relative maximum. Because this rarely happened, in all the other cases we kept the number of groups that were associated with a decrease in slope in the *Pseudo-F* profile. The application of *K-means* and the calculation of the *pseudo-F* were performed with the multivariate analysis module GINKGO (De Cáceres et al. 2007) in the VEGANA package (De Cáceres et al. 2003). The profile curves of the values of the *pseudo-F* for the 12 continuous characters included in the data matrix are provided (SUPPLEMENTARY FIG. 1).

*DNA isolation and sequencing.*—Genomic DNA was obtained from fresh samples and herbarium specimens (voucher information is listed in SUPPLEMENTARY TABLE I; for more detailed locality information see Gaya 2009) and isolated with the Puregene Kit (GENTRA Systems, Minneapolis, Minnesota) following the manufacturer's protocol for filamentous fungi. DNA concentration was determined by visual comparison with positive control (λ 100 ladder, concentration 10, 20, 40 ng) on an ethidium-bromide-stained TBE agarose gel. Symmetric polymerase chain reactions (PCR) were prepared for a 50.0 μL final volume containing 31.7 μL sterile double-distilled water, 5.0 μL 10 × Taq polymerase reaction buffer (Boehringer-Mannheim, Indianapolis, Indiana), 5.0 μL 2.5 mM dNTPs, 0.3 μL Taq DNA polymerase (Boehringer-Mannheim), 2.5 μL for each of the 10 μM primers ITS1F, ITS1 or ITS5 with ITS4 (White et al. 1990, Gardes and Bruns 1993), 1.5 μL 10 mg/mL bovine serum albumin (BSA; BioLabs), 0.5 μL of 50 mM $MgCl_2$ and 1 μL of template genomic DNA. PCR was performed on PerkinElmer GeneAmp 2400 under these conditions: one cycle of 1 min at 95 C linked to 40 cycles of 1 min at 95 C, 45 s at 52 C, and 2 min at 72 C with the last step increased by increments of 5 s for the last 15 cycles. A final extension step of 10 min at 72 C was added, after which the samples were kept at 4 C. The PCR products were purified with either the Cycle-Pure Kit (E.Z.N.A.) or low binding regenerated cellulose 30 000 NMWL (nominal molecular weight limit) filter units (Millipore, Billerica), following manufacturer instructions. Both strands of the purified PCR products were sequenced with PCR primers used for the symmetric amplification and additional primers 5.8S and 5.8SR (Vilgalys and Hester 1990) or ITS2 and ITS3 (White et al. 1990). Sequencing reactions were prepared in 10 μL final volume with BigDye Terminator 3.1 (ABI PRISM, PerkinElmer Biosystems, Wellesley, Massachusetts) and following the manufacturer's instructions. Sequenced products were precipitated with 26 μL deionized sterile water and 64 μL 95% ethanol before they were loaded on an ABI Prism 3730 automated DNA sequencer (PerkinElmer, Applied Biosystems).

*Sequence alignment.*—Sequence fragments were subjected to BLAST queries for an initial verification of their identities. They were assembled with Sequencher 4.1 (Gene Codes Corp., Ann Arbor, Michigan) and Sequencer

Navigator 1.0.1 (Applied Biosystems Inc. 1989–1994) and aligned manually with MacClade 4.01 (Maddison and Maddison 2001). The delimitation of ambiguous regions was done manually following Lutzoni et al. (2000). Their unequivocal coding and the elaboration of symmetric step matrices for each of these coded characters were generated with the program INAASE 2.3b (Lutzoni et al. 2000). All DNA sequences have been deposited in GenBank (see SUPPLEMENTARY TABLE I for GenBank accession numbers), and the alignment is available in TreeBASE (accession number S10717).

*Phylogenetic analyses.*—Phylogenetic analyses were performed with maximum parsimony (MP) and maximum likelihood (ML) optimization criteria as implemented in PAUP* 4.0b10 for UNIX and Macintosh (Swofford 2002), except for the ML bootstrap analysis, performed with GARLI 0.96 (Zwickl 2006), and Bayesian analysis with BAli-Phy (Suchard and Redelings 2006) and MrBayes 3.0b4 (Huelsenbeck and Ronquist 2001). Maximum parsimony searches were implemented on the separate morphology dataset (MP1), on three ITS datasets (MP2, 60 sequences; MP3 and MP4, 68 sequences) and on the morphological dataset combined with the ITS sequences (MP5) for the same 60 specimens included in MP1 and MP2. Maximum likelihood and Bayesian analyses were performed only on the 68-sequence matrix. For the separate set of analyses on morphology (MP1) all changes among character states were equally weighted and unordered. To simultaneously accommodate taxa with multiple character states resulting from uncertainty or polymorphy, the VARIABLE option in PAUP* was used (SUPPLEMENTARY MATERIALS AND METHODS).

In the MP analyses of the ITS datasets (MP2, MP3 and MP4) and the combined dataset (MP5) gaps were used as a fifth character state and symmetric step matrices were created for unambiguously aligned portions of the alignments as follows: the options SHOW CHARACTER STATUS/FULL DETAILS/HIDE EXCLUDED CHARACTERS from the data menu in PAUP* were selected. From the resulting table the States column showing all nucleotide states found at each of the unambiguously aligned variable sites was saved as a separate text file. This file then was used as input for STMatrix 2.1 (François Lutzoni and Stefan Zoller, Dept. Biology, Duke Univ.; http://www.lutzonilab.net/downloads), which generates a step matrix (in Nexus format) by calculating frequencies of reciprocal changes from one state to another and converting them into costs of changes using the negative natural logarithm of the frequencies (Felsenstein 1981, Wheeler 1990). ITS1, ITS2 and 5.8S each were subjected to a specific symmetric step matrix. Constant sites and ambiguously aligned regions were removed from MP searches and MP bootstrap analyses.

Phylogenetic signal from ambiguously aligned regions was recovered without violating positional homology with the program INAASE 2.3b. (Lutzoni et al. 2000). In MP2 and MP4 both unambiguously aligned sites and INAASE characters were included while MP3 was restricted to unambiguously aligned sites. The combined data matrix (MP5) included the morphological data matrix, the coded INAASE characters and the ITS data matrix used in the MP2 analysis.

All MP searches were performed with heuristic searches with TBR (tree bisection-reconnection) branch swapping, MULTREES option in effect, and collapsing branches with maximum branch length equal to zero. For the separate analyses on morphology (MP1) an exploratory search first was performed with 1000 random addition sequences (RAS). To improve the resolution and reduce computing time we performed a search in two steps. In the first step we estimated a lower boundary for the length of the most parsimonious trees, with 1000 RAS and saving only one tree per replicate. In the second step we searched for all equally most parsimonious trees, with 10 000 RAS, saving all trees only when swapping in a tree equal to or shorter than the shortest tree found in the first step. The remaining MP analyses were conducted with 1000 random addition sequences (RAS) in one step by saving all trees as soon as TBR swapping was initiated.

Maximum likelihood searches on the ITS dataset of 68 sequences included constant sites. Selection of the nucleotide substitution model and the estimation of most parameters was done with a hierarchical likelihood ratio test (HLRT) (Huelsenbeck and Crandall 1997), with a bottom-up approximation and with the Akaike information criteria (AIC) (Akaike 1973) using the program Modeltest 3.06 (Posada and Crandall 1998). Simultaneously HLRT also was calculated in PAUP* with a top-down approximation to test whether alternative models were obtained. In this study the aim of estimating the model with different methods was to empirically test the sensitivity of the ML analyses to the differences in the substitution model selected by one or another approximation, to demonstrate whether different models converge on a same topology.

Three ML searches were implemented with each of the models obtained with the methods detailed above. For HLRT (ML1) the selected substitution model was TrN+G (Tamura and Nei 1993), for AIC (ML2) was GTR+I+G (Lanave et al. 1984), and for HLRT calculated in PAUP* with a top-down approximation (ML3) the selected model was HKY+G (Hasegawa-Kishino-Yano 1985; see SUPPLEMENTARY MATERIALS AND METHODS for details on the model parameters). In the three cases ML heuristic searches were executed with the same dataset as for MP3, with 1000 RAS, number of rearrangements limited to 40 000/RAS, TBR branch swapping, MULTREES option in effect, saving all trees and collapsing branches with maximum branch length equal to zero.

Branch support for MP and ML was assessed by bootstrap analyses (Felsenstein 1985) with full heuristic searches. The number of RAS per bootstrap replicate was calculated taking into consideration the frequency at which the shortest or most likely tree was found during the heuristic search with the original dataset. In MP1, MP2 and MP5 10 000 parsimony bootstrap replicates were performed, using four RAS per bootstrap replicate and by saving no more than 10 trees per bootstrap replicate, with the same parameters as in the initial MP analyses. For MP3 1000 bootstrap replicates were performed with two RAS per bootstrap replicate, with the same parameters as for the initial MP analyses. In MP4 300 bootstrap replicates were performed, using two RAS per bootstrap replicate with the same parameters as for the initial MP4 analysis. For ML

bootstrap analysis 2000 bootstrap replicates were performed with two RAS (2 searchreps) in Garli 0.96 (Zwickl 2006) under the same substitution models selected for the ML searches.

We also used Bayesian posterior probabilities (PP), approximated by sampling trees with a Markov chain Monte Carlo with Metropolis coupling method as implemented in MrBayes 3.0b4 (Huelsenbeck and Ronquist 2001) to assess confidence for relationships revealed by ML searches. Bayesian analyses were initiated on random trees and run 5 000 000 generations, sampling the Markov chains every 100 generations. To ensure that all trees from the burn-in were excluded the majority rule consensus tree was calculated with PAUP* using only the last 37 500 of the 50 000 trees sampled. The exclusion was made by plotting the log-likelihood values against the time generation and noting when log-likelihood values reached a stable equilibrium value (Huelsenbeck and Ronquist 2001). We repeated the Bayesian tree sampling as described above to confirm the result from the first Bayesian analysis. Both runs converged on nearly the same average likelihood score −1620.31 and −1620.23. The final majority rule consensus tree and nodal posterior probabilities were obtained by pooling all trees selected from both runs (75 000 trees).

In addition, we conducted another two Bayesian analyses with the software package BAli-Phy (Suchard and Redelings 2006) on the ITS dataset of 68 sequences, divided into three partitions, ITS1, 5.8S and ITS2. In the first analysis (B1) the alignment first was censored to remove ambiguous regions and held constant for all partitions. We then estimated the phylogeny and model parameters conditional on a fixed alignment. In the second analysis (B2) the alignment was not censored but instead sampled from the joint posterior distribution of phylogeny, alignment of ITS1 and ITS2 and model parameters (Redelings and Suchard 2005). In this case the alignment of the 5.8S partition was fixed because the number of accepted mutations in that region was sufficiently low to render the alignment essentially unambiguous. This approach takes into account alignment uncertainty by integrating overall alignments in proportion to their posterior probability. Both analyses were conducted with the Markov chain Monte Carlo (MCMC) algorithm.

In B2 we let all three partitions share a single topology and placed a uniform prior distribution on the topology. To constrain the two spacers (ITS1, ITS2) to have the same rate, but allow a different rate for the 5.8S, we introduced the parameter $\mu^{1,3}$ for the mean branch length in the first and third partitions and $\mu^2$ for the mean branch length in the second partition. Also we let branch lengths in each partition be obtained by scaling a set of shared branch lengths $\mathbf{T}$ by either $\mu^{1,3}$ or $\mu^2$, depending on the partition. We placed an exponential prior distribution with mean = 1 on the length $T^b$ for each branch b. As a result the prior distribution on each branch length in partition $p$ is an exponential distribution with mean $\mu^p$. In addition we placed an exponential prior with mean = 1 on the hyperparameters $\mu^{1,3}$ and $\mu^2$ representing the mean branch length in each partition (Suchard et al. 2003).

For both analyses (B1 and B2) the same model of nucleotide substitution, TrN (Tamura and Nei 1993), was implemented. We used the +gwF formulation (Goldman and Whelan 2002) of the TrN model instead of the original +F formulation (Cao et al. 1994) because it introduces an additional parameter $f$ an element of [0, 1] that specifies whether common nucleotides occur more frequently because they are more highly conserved (low $f$) or because they are more often proposed as replacements (high $f$). In our analysis the first and third partitions were constrained to have the same substitution model parameters. In addition we allowed substitution rate heterogeneity in the first and third partitions in a manner described below and did not estimate substitution rate heterogeneity in the model for the second partition (5.8S).

Our model of substitution rate heterogeneity assumes that the substitution rates at each site are one or three separate rates $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ with prior probabilities $\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$. As is customary we constrained the average substitution rate to 1 by rescaling the computed rate matrices. The substitution model therefore was characterized fully by the parameters $\kappa^{1,3}_{AG}$, $\kappa^{1,3}_{TC}$, $\boldsymbol{\pi}^{1,3}$, $f^{1,3}$, $\mathbf{r}$, $\mathbf{q}$, $\kappa^2_{AG}$, $\kappa^2_{TC}$, $\boldsymbol{\pi}^2$, $f^2$, with a total of 16 degrees of freedom. We placed a log-Laplace prior with mode 2 and scale 0.25 on the $\kappa_{PUR}$ and $\kappa_{PYR}$ parameters, a uniform Dirichlet prior on $\boldsymbol{\pi}^{1,3}$ and $\boldsymbol{\pi}^2$, a uniform (0,1) prior on the $f$ parameters, a Dirichlet (2.5, 2.5, 2.5) prior distribution on $\mathbf{q}$, and a Dirichlet (2, 2, 2) prior distribution on $\mathbf{r}$.

In the B2 analysis the RS07 indel model from Redelings and Suchard (2007) was used for the alignments in the first and third partitions. These partitions share a single insertion/deletion rate $\lambda$ relative to the substitution rate and a single gap extension probability $\varepsilon$. Thus the insertion/deletion model parameters ($\lambda$, $\varepsilon$) add two degrees of freedom. We note that the insertion/deletion rate given here is for insertion or deletion separately; the combined rate would be $2\lambda$.

Parameters for the B1 and B2 analyses then were estimated by combining results from several independent MCMC runs (Beiko et al. 2006). In B1 analyses two independent runs were performed, whereas in B2 six independent runs were executed. In both cases 100 000 iterations were run, sampling the Markov chains every 10 iterations. The first 10 000 samples of each run were discarded as burn-in, the exclusion was made as described above, retaining a total of 9000 sampled trees per run. The final majority rule consensus tree was obtained by pooling all trees selected from all runs after burn-in.

To determine whether our MCMC estimates are reliable we sought to show that the burn-in period is sufficient to ensure stabilization of parameter values and that a sufficiently large number of samples were collected following convergence to ensure precise estimates of the posterior probabilities for each split. We focused primarily on the B2 analysis because of the difficulty in designing MCMC algorithms that converge when sampling multiple sequence alignments.

We determined that the tree topology and all numerical parameters stabilized before iteration 1320 in each of the six runs of the Markov chain (see SUPPLEMENTARY MATERIALS AND METHODS). Hence we concluded that removing 10 000 iterations was sufficiently conservative to ensure stabilization of parameter values. We also computed potential scale

reduction factors (PSRF) to assess convergence of numerical parameters after stabilization. In each case the PSRF was less than 1.002 (see SUPPLEMENTARY MATERIALS AND METHODS), and so no lack of convergence was detected.

We also compared estimates of the equilibrium distribution of topologies for each of the six runs in the B2 analysis with the Robinson-Foulds (1981) distance. We obtained 600 topology samples from each of the six runs by subsampling topologies from the last 50 000 iterations of each run. In each case the average distance between trees taken from the same run was 20.44–20.69, while the average distance between two trees taken from different runs varied 20.40–20.68, indicating that topologies sampled from different runs had not stabilized in different regions of tree space. Visual comparison of the samples from different runs via multidimensional scaling yielded strong support for this conclusion (Hillis et al. 2005).

We compared estimates for the posterior probability of splits across runs to assess convergence. Following Beiko et al. (2006), we graphed the range of split frequency estimates across the six different runs of the B2 analysis to show that different starting values produced similar estimates (FIG. 1); we also depicted point estimates of the split frequencies from the B1 analysis for comparison. We considered splits that obtained a posterior probability of 0.1 or higher in any run of the B1 or B2 analysis. The relatively short length of the vertical bars indicates that the variance of the estimates across runs was relatively low and that different runs were not trapped in separate modes.

To determine the precision of our estimates of split frequencies we first computed the standard deviation of split frequencies for each split. We then averaged these values across splits to obtain the average standard deviation of split frequencies (ASDSF) as reported in MrBayes (Huelsenbeck and Ronquist 2001). For the B1 analysis the average value (ASDSF) was 0.003 and the maximum 0.011. For the B2 analysis the average value was 0.004 and the maximum 0.014. The ASDSF values indicate that the split frequencies were estimated with an extremely high degree of precision for most splits; the maximum values indicate that even the lowest degree of precision is acceptably high.

*Congruence assessment between morphology and ITS data.*—Before combining the morphological and ITS datasets for the MP5 analysis topological congruence among partitions was assessed with a reciprocal 70% BP threshold (Mason-Gamer and Kellogg 1996). Topological conflicts accordingly were assumed to be significant if two different relationships (one monophyletic for one data partition and the other non-monophyletic for the other data partition) for the same set of taxa were both supported with bootstrap values ≥ 70% by the morphological versus the ITS datasets.

## RESULTS

*Morphological characters selected.*—Most of the 38 phenotypic characters included in the data matrix are those currently used in the determination and description of species of genus *Caloplaca*, especially for the lobed taxa (for a broader discussion on the morphological characters of the *C. saxicola* group see Gaya 2009). The 57 anatomical, morphological and ecological characters considered potentially useful for the phylogenetic analysis are available online (SUPPLEMENTARY TABLE II).

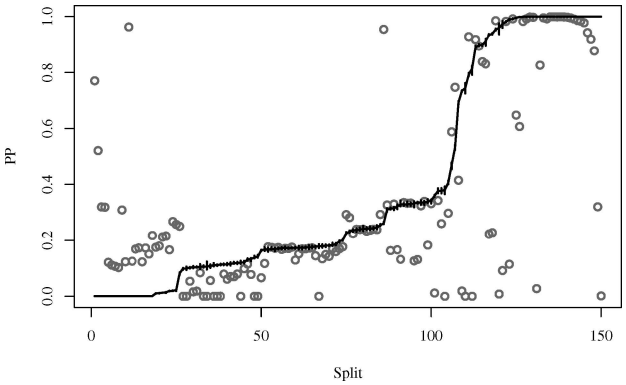


FIG. 1. Convergence diagnostics and differences between B1 and B2 analyses. The black line shows the posterior probability (y axis) for each split (x axis) in the B2 analysis. Gray circles indicate the posterior probability of each split in the B1 analysis. Each integer on the x axis represents one of the splits that obtained a posterior probability of 0.1 or greater in any MCMC run for either analysis. The splits then were sorted in order of increasing posterior probability in the B2 analysis. Vertical bars represent the range of variation across runs for estimates of the posterior probability of the same split.

*Phylogenetic analyses of morphological characters.*—All 38 phenotypic characters used in the MP1 analysis were parsimony informative (TABLE I). This search yielded 71 equally most parsimonious trees (TABLE II) of 618 steps, summarized in the strict consensus tree (FIG. 2A) (consistency index [CI] = 0.741, retention index [RI] = 0.669). Most relationships are unresolved and only eight internodes had BP > 50%, mostly due to the low number of morphological characters (FIG. 2A).

*Phylogenetic analyses of ITS.*—The final data matrix of 60 sequences included 550 sites (TABLE I). In MP2 a total of 22 ambiguously aligned regions were delimited, resulting in the exclusion of 188 nucleotide sites. Moreover, 278 constant sites were removed for a total of 466 sites excluded. The 84 remaining characters were combined with 22 INAASE-coded characters, for a total of 106 variable characters; of those 77 were parsimony informative.

The size of the final data matrix of 68 sequences consisted of 551 sites (TABLE I), from which a total of 22 ambiguously aligned regions (189 sites) were excluded from MP3, MP4, ML1, ML2, ML3 and Bayesian analyses (MrBayes and BAli-Phy B1). Of the 362 characters subjected to ML and Bayesian analyses

TABLE I. Synopsis of datasets used in MP, ML and Bayesian analyses

| | 60-Morph (MP1) | 60-ITS (MP2) | 60-Morph + ITS (MP5) | 68-ITS (MP3) | 68-ITS (MP4) | 68-ITS (ML1, 2, 3 & Bayesian) | 68-ITS B1 | 68-ITS B2 |
|---|---|---|---|---|---|---|---|---|
| Total number of characters (+ INAASE characters) | 38 | 550 (+22) | 588 (+22) | 551 | 551 (+20) | 551 | 551 | 551[b] 570 (566, 575) |
| Number of excluded sites[a] | NA | 466 | 466 | 463 | 463 | 189 | 189 | 0 |
| Number of constant sites | NA | 278 | 278 | 274 | 274 | 274 | 274 | 355[b] 381 (375, 389) |
| Number of aligned variable sites | NA | 84 | 84 | 88 | 88 | 88 | 88 | 196[b] 188 (184, 192) |
| Number of ambiguously aligned regions (sites) | NA | 22 (188) | 22 (188) | 22 (189) | 22 (189) | 22 (189) | 22 (189) | 0 |
| Number of ambiguous regions recoded with INAASE | NA | 22 | 22 | 0 | 20 | 0 | 0 | 0 |
| Total number of analyzed characters | 38 | 106 | 144 | 88 | 108 | 362 | 362 | 551[b] 570 (566, 575) |
| Number of parsimony informative characters | 38 | 77 | 115 | 59 | 79 | NA | NA | NA |

[a] Excluded sites include ambiguously aligned regions.

[b] For the B2 analysis only we provide the posterior median and 95% credible interval next to the value for the initial alignment, number of constant sites and variable sites, where appropriate. NA = Not applicable.

TABLE II. Synopsis of analytical results and internode support

| Analyses type | 60-Morph (MP1-BP) | 60-ITS (MP2-BP) UNAMB + I | 60-Morph + ITS (MP5-BP) UNAMB + I | 68-ITS (MP3-BP) UNAMB | 68-ITS (MP4-BP) UNAMB + I | 68-ITS (ML1-BP) UNAMB | Bayesian-PP UNAMB | 68-ITS B1-PP UNAMB | 68-ITS B2-PP UNAMB + AMB |
|---|---|---|---|---|---|---|---|---|---|
| Number of resolved internodes | 27 | 41 | 47 | 30 + 4* = 34 | 37 + 6* = 43 | 32 + 5* = 37 | 31 + 5* = 36 | 32 + 5* = 37 | 39 + 5* = 44 |
| Number of significantly supported internodes | 4 | 26 | 27 | 9 + 4* = 13 | 24 + 5* = 29 | 10 + 5* = 15 | 18 + 4* = 22 | 17 + 5* = 22 | 27 +5 * = 32 |
| Number of equally most parsimonious or likely trees | 71 | 48 | 192 | 393 | 1299 | 7 | NA | NA | NA |

Internodes were considered significantly supported when bootstrap values were ≥ 70% and posterior probabilities were ≥ 95%. BP, bootstrap proportions; I, addition of INAASE characters; ML, maximum likelihood analyses; MP, maximum parsimony analyses; PP, posterior probabilities; B1, Bayesian analysis implemented in BAli-Phy with fixed alignment; B2, Bayesian analysis implemented in BAli-Phy with estimated alignment; UNAMB, unambiguously aligned sites; AMB, ambiguously aligned sites. To compare the topologies derived from the dataset with 68 specimens to the topologies obtained with the dataset of 60 specimens, internodes supporting additional taxa in the dataset of 68 specimens are indicated separately and denoted by an asterisk (*).

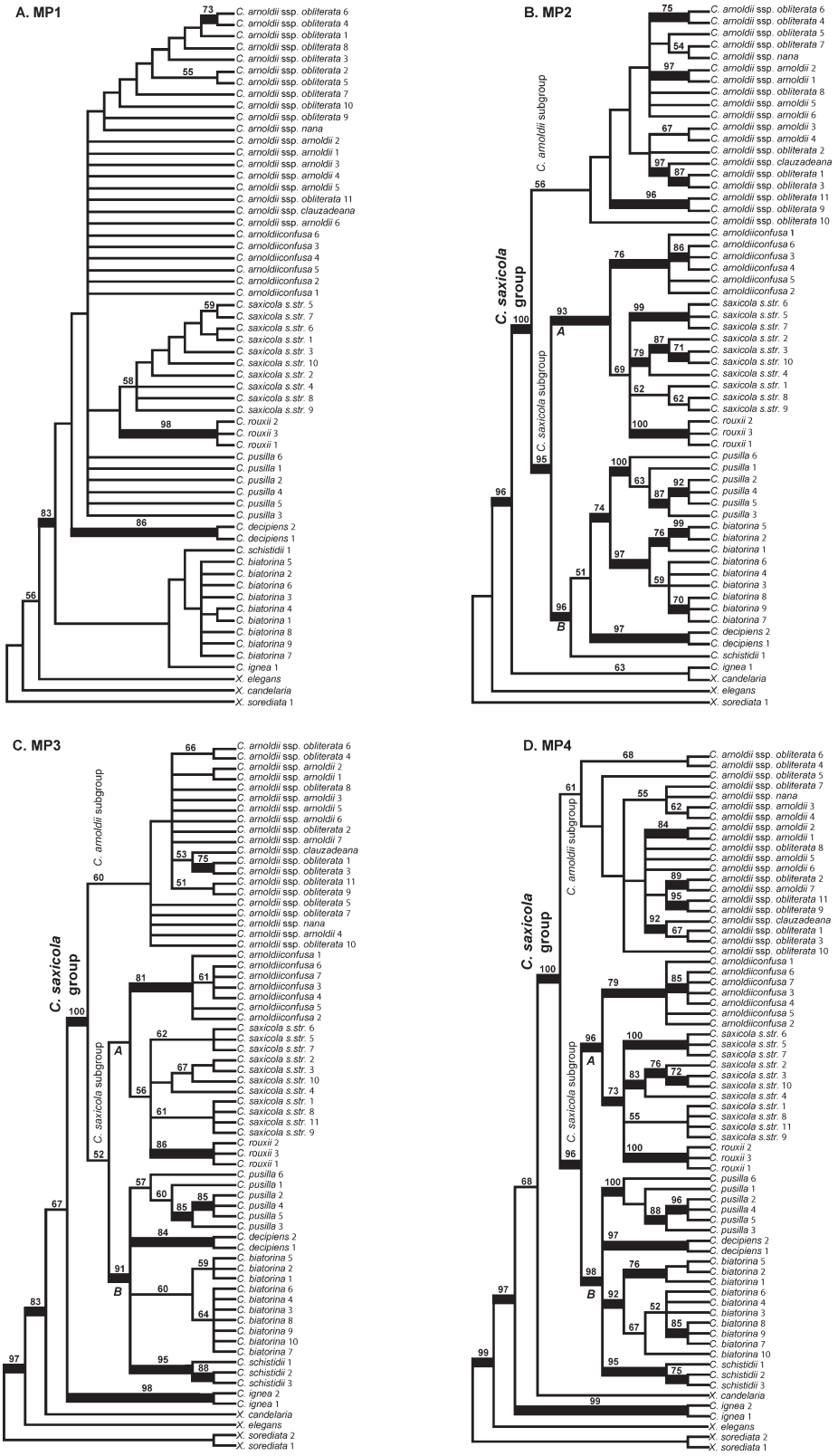


FIG. 2. A. MP1 analysis. Phylogenetic relationships among 56 individuals from the *C. saxicola* group as revealed by maximum parsimony analyses of the morphological dataset (38 characters by 60 OTUs) and summarized by a strict consensus tree of 71 equally most parsimonious trees. B. MP2 analysis. Phylogenetic relationships among 56 individuals from the *C. saxicola* group as revealed by maximum parsimony analyses of the ITS dataset including 22 INAASE characters (106 characters

274 were constant and 88 were variable. In MP3 and MP4 the 274 constant sites were removed for a total of 463 sites excluded. In MP3 from the 88 remaining sites 59 were parsimony informative. In MP4 20 INAASE characters were combined with the 88 characters for a total of 108 variable characters, and of those 79 were parsimony informative.

The unequally weighted MP2 (unambiguously aligned regions and 22 INAASE characters) search yielded 48 equally most parsimonious trees (TABLE II) of 513.39 steps, which were part of one island that was hit 983 times out of 1000 RAS (CI, excluding uninformative characters = 0.7086; RI = 0.877) (FIG. 2B). The unequally weighted MP3 (unambiguously aligned regions) search generated 393 equally most parsimonious trees of 239.74 steps, which were part of six islands, one of which was hit 990 times out of 1000 RAS, whereas the remaining five were hit only once (CI, excluding uninformative characters = 0.588; RI = 0.874) (FIG. 2C). The unequally weighted MP4 (unambiguously aligned regions and 20 INAASE characters) search yielded a total of 1299 equally most parsimonious trees distributed into five islands, one of which was hit 996 times out of 1000 RAS, and the remaining four each hit once. The length of the best trees was 493.41 steps (CI, excluding uninformative characters = 0.678; RI = 0.887) (FIG. 2D).

The three MP analyses on the ITS dataset generated similar topologies, but with discrepancies in the resolution and the level of confidence. In adding the eight sequences from GenBank, representing minor genetic variations within species already included, we greatly increased the number of equally most parsimonious trees (48 in MP2 vs. 1299 in MP4). Regarding the number of well supported internodes (BP $\geq$ 70%), MP2 and MP4 analyses showed similar numbers (26 in MP2 and 29 in MP4, 21 of which were shared). In general none of the discrepancies among topological bipartitions involved groupings with bootstrap values $\geq$ 70%.

Adding INAASE characters reduced the apparent level of resolution in the MP4 versus the MP3 search because branches with lengths equal to zero were collapsed during the search (1299 in MP4 vs. 393 trees in MP3). Regarding the total number of resolved internodes, MP4 showed superior results compared to MP3 (43 vs. 34). These differences occurred mainly in the *C. arnoldii* subgroup and, to a lesser extent, in the *C. biatorina* clade (FIG. 2C, D). As expected confidence was higher in MP4 than MP3 (29 internodes BP $\geq$ 70% vs. 13 in MP3). There were five internodes with BP $\geq$ 70% in MP4 where support $\geq$ 50% was not obtained in MP3 but only one case where MP3 significantly supported a node that MP4 did not.

The ML1 analysis (model selected with HLRT ModelTest) generated seven equally most likely trees ($-\ln$ = 1467.34455) that were part of five islands, obtained 774 times out of 1000 RAS. In ML2 (model selected with AIC) seven equally most likely trees also were found ($-\ln$ = 1460.542445), part of five islands obtained 694 times out of 1000 RAS. Finally, ML3 (top-down HLRT) also generated seven equally most likely trees ($-\ln$ = 1469.77750), part of five islands obtained 752 times out of 1000 RAS. ML1 (TrN+G) and ML2 (GTR+I+G) trees were exactly the same, whereas for ML3 (HKY+G) six of the seven trees also were identical; the seventh tree showed a slightly different topology in the *C. arnoldii* clade. These results indicate that the data were not sensitive to the different models selected for ML1 and ML2, whereas some sensitivity was observed when selecting the simplest model obtained with the top-down approach. Given the negligible effect between the most complex model and the TrN+G model, we selected the model with the least number of parameters (TrN+G), to improve computational efficiency for the ML1 search. Henceforth, the ML1 analysis/search will refer exclusively to the ML1 search with the TrN+G model.

The ML1 search showed a loss in resolution over most MP analyses (FIG. 3A). However, when the MP and ML analyses were restricted to unambiguously aligned ITS sites (MP3 and ML1), that is identical data matrices, the ML analysis was superior to the MP analysis both in numbers of resolved internodes and significantly supported internodes (TABLE II). When adding INAASE characters to the ITS data matrix, which can be analyzed with MP but not by ML in its current implementations, MP analyses generated more resolved trees with higher confidence than ML analyses, resulting in 43 internodes resolved by MP4 versus 37 by ML1. For 15 internodes significantly supported in MP4 the ML1 analysis resolved eight

---

←

by 60 OTUs) and summarized by a strict consensus tree of 48 equally most parsimonious trees. C. MP3 analysis. Phylogenetic relationships among 62 individuals from the *C. saxicola* group as revealed by maximum parsimony analyses of the ITS dataset without INAASE characters (88 characters by 68 OTUs) and summarized by a strict consensus tree of 393 equally most parsimonious trees. D. MP4 analysis. Phylogenetic relationships among 62 individuals from the *C. saxicola* group as revealed by maximum parsimony analyses of the ITS dataset including 20 INAASE characters (108 characters by 68 OTUs) and summarized by a strict consensus tree of 1299 equally most parsimonious trees. Bootstrap support $\geq$ 50% is shown above branches. Support values $\geq$ 70% are highlighted by thicker branches.

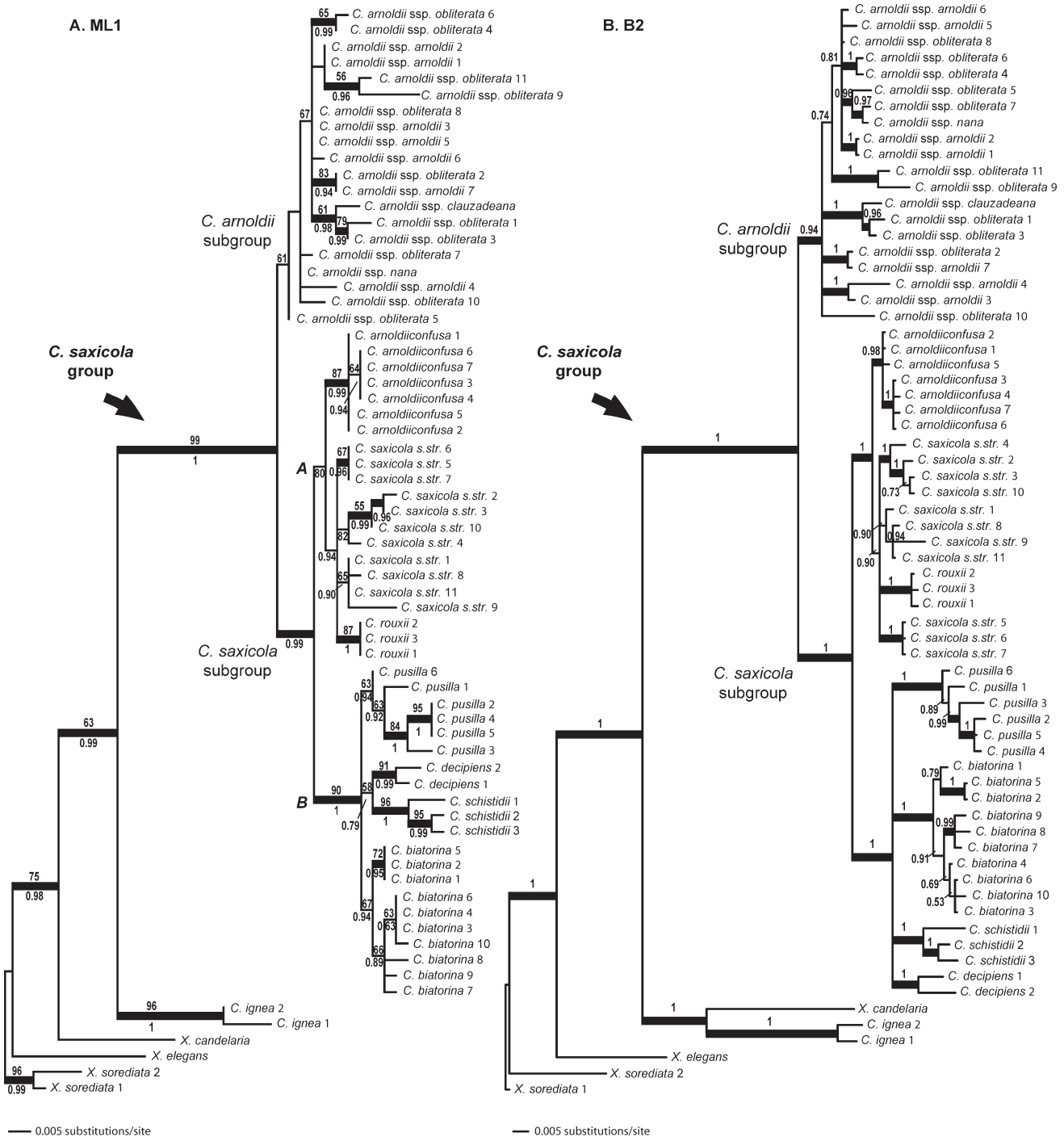


FIG. 3. A. ML1 analysis. Phylogenetic relationships among 62 individuals from the *C. saxicola* group as revealed by maximum likelihood analyses of the ITS dataset (362 characters by 68 OTUs). One of the seven equally most likely trees (−ln likelihood = 1467.34455). Numbers above each internode represent bootstrap values ≥ 50%. Numbers below internodes indicate posterior probabilities ≥ 0.5. B. B2 analysis. Phylogenetic relationships among 62 individuals from the *C. saxicola* group as revealed by the Bayesian analysis of nonaligned sequences of ITS1 and ITS2 and aligned sequences of 5.8S (BAli-Phy) of the ITS dataset (551 characters by 68 OTUs), and summarized by a majority consensus tree. Posterior probabilities ≥ 0.5 are shown above branches. Internodes with BP values ≥ 70% or PP ≥ 0.95 are highlighted by thicker lines.

with BP < 70%, and seven without support (i.e. below 50%).

With regard to the Bayesian analysis performed with MrBayes the number of significantly supported internodes, 22, (FIG. 3A) surpassed MP3 and ML1. However, for 10 internodes significantly supported in MP4 the Bayesian analysis recovered seven with PP < 0.95, and three without support. Only two internodes with BP < 70% in MP4 received a PP ≥ 0.95 and in these two cases significant support was also recovered in MP3. Therefore, there were no significantly supported internodes in the ML1 and Bayesian analyses performed with MrBayes that did not recover significant bootstrap values in any of the parsimony searches. Regarding the number of resolved internodes in the Bayesian analysis, these are surpassed by MP4 (43 vs. 36), and ML1 (37 vs. 36). Finally, the Bayesian analysis showed a single significant discrepancy among topological bipartitions with MP4 (PP = 0.96 and BP = 72%), which is shared by MP3 but without support.

Regarding the Bayesian analyses performed with BAli-Phy, we describe parameter estimates from the B2 analysis. Estimates from the B2 analysis are preferred because it did not rely on a priori alignment of ITS1 and 2, made use of information in shared indels and used more of the ITS data. However, parameter estimates from the B1 analysis are similar, with the exception that the B1 model does not include the parameters λ and ε. The substitution rate $\mu^{1,3}$ for the ITS spacers 1 and 2 (corresponding to partitions 1 and 3) was about 13.9 times higher than the substitution rate $\mu^{2}$ for 5.8S itself. We did not estimate an indel rate for 5.8S; the ratio λ between the indel rate and the substitution rate for the ITS spacers was about 0.1004, which is nearly double the values obtained by Cartwright (2009) from mammalian introns and by Lunter (2007) from entire mammalian genomes. However, this difference could reflect the fact that the RS07 model assumes that insertion/deletion lengths follow a geometric distribution, which makes unit-length indels a priori less likely compared to a power law used by Cartwright and the mixture of geometric distributions used by Lunter. The mean length of insertions/deletions was about 1.2, indicating that most insertions and deletions affected only a single nucleotide. The ITS spacers had different nucleotide frequencies $\boldsymbol{\pi}^{1,3}$ than 5.8S $\boldsymbol{\pi}^{2}$: the spacers had a GC content of about 62% compared to a GC content of 49% for 5.8S. Furthermore, the variation in the rates **r** indicates that some sites in the spacers are accumulating substitutions at a rate that is 12 times slower than some other sites. Thus some nucleotides in the spacers may be under purifying selection (SUPPLEMENTARY TABLE III).

The number of significantly supported internodes in B1 and B2 analyses was respectively 22 and 32 (TABLE II). With the exception of B2 the analysis B1 (tree not shown) was surpassed only by MP4, where eight internodes significantly supported in MP4 showed PP < 0.95 in B1, even though three internodes with PP ≥ 0.95 in B1 did not recover significant support in MP4. Unsurprisingly, the number of supported internodes in B1 was the same as in the similar Bayesian analyses implemented in MrBayes. In general confidence for the *C. biatorina* clade was lower for the B1 analysis compared to the other analyses. However, the number of resolved internodes (37) exceeded MP3 and the other Bayesian analysis.

Analysis B2 (FIG. 3B) showed a higher performance, surpassing all other analyses. Seven internodes significantly supported in B2 showed BP < 70% in MP4, 18 internodes highly supported in B2 recovered BP < 70% in ML1 and 12 internodes with PP ≥ 0.95 in B2 were not significantly recovered in the Bayesian approach with MrBayes. Regarding the total number of resolved internodes, again B2 showed superior results (44) compared to the previous analyses.

Therefore, the additional signal recovered from ambiguously aligned regions without violating positional homology with INAASE did not surpass the Bayesian approach when alignment estimation was part of the phylogenetic search.

*Phylogenetic analyses of combined morphological and molecular data (MP5).*—Topologies resulting from the morphological and molecular datasets when analyzed separately showed certain differences (MP1 vs. MP2). However, none of these differences were reciprocally supported with bootstrap proportions ≥ 70% (Mason-Gamer and Kellogg 1996). Therefore, these differences were not significant and the two datasets were combined. The size of the combined data matrix was 610 characters (TABLE I). As in MP2 a total of 188 sites, representing 22 ambiguously aligned regions, were excluded together with 278 constant sites. The 84 remaining sites were combined with 22 coded INAASE characters and 38 phenotypic/ecological characters for a total of 144 variable characters, 115 of which were parsimony informative. The unequally weighted MP5 search yielded 192 equally most parsimonious trees (TABLE II) of 1180.10 steps. These trees were part of five islands, one of which was hit 405 times out of 1000 RAS, whereas the rest were hit only once (CI, excluding uninformative characters = 0.736; RI = 0.818) (FIG. 4).

The MP5 search showed an apparent increase in the number of equally most parsimonious trees

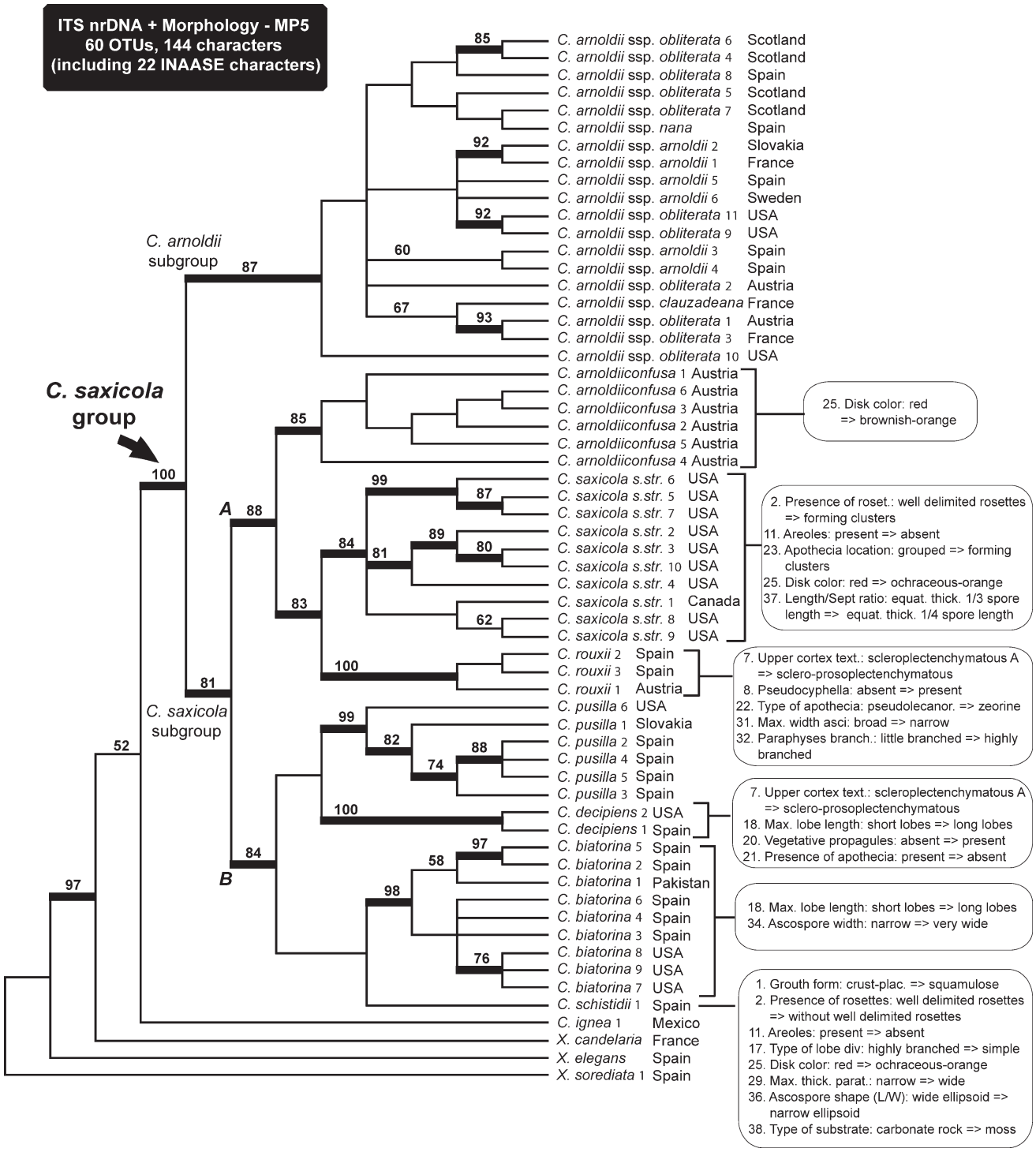


FIG. 4. MP5 analysis. Phylogenetic relationships among 56 members of the *C. saxicola* group as revealed by a maximum parsimony search on the combined morphology and ITS datasets including 22 INAASE characters and summarized by a strict consensus tree of 192 equally most parsimonious trees. Bootstrap support ≥ 50% is shown above branches. Support values ≥ 70% are highlighted by thicker branches. Boxes to the right of the tree provide sets of unequivocal synapomorphies supporting species as defined by Gaya (2009). Geographic origin of specimens is indicated after each taxon name.

compared to MP2 (48 vs. 192 trees), but as for MP3 and MP4 this is due to the fact that branch lengths equal to zero are collapsed and these equally parsimonious trees are not counted. The number of internodes is a better proxy for resolution. Using a strict consensus to summarize all equally parsimonious trees, the combined analysis (MP5) resolved six additional internodes over the MP2 analysis restricted to the ITS, and 20 additional internodes over the MP1 analysis of the morphological/ecological dataset. The MP5 analysis resulted in 27 internodes with BP $\geq$ 70%, one more than MP2 and 23 more than MP1. It also surpassed the 68 individual datasets restricted to ITS data even when considering the internodes relating to the eight added GenBank sequences, except MP4 and B2. MP5 and B2 analyses, showed the same level of confidence, and both searches recovered extremely similar topologies with comparable numbers of resolved internodes (47 in MP5 vs. 44 in B2). Moreover, in some clades, as in *C. saxicola* s. str., MP5 confidently recovered important relationships that were not resolved in B2. Therefore, we considered the phylogeny derived from MP5 as our best estimate of relationships within the *Caloplaca saxicola* group.

*Phylogenetic relationships.*—The *C. saxicola* group (sensu Gaya et al. 2008) was recovered as monophyletic (BP = 100%), consisting of two main lineages (*C. arnoldii* and *C. saxicola* subgroups, FIG. 4). The *C. arnoldii* subgroup corresponded to the *C. arnoldii* complex (Gaya 2009) with the remaining species within the *C. saxicola* group placed within the *C. saxicola* subgroup. The *C. arnoldii* subgroup was significantly supported by the MP5 analysis (FIG. 4), but there is still a large degree of uncertainty with regard to delimitation of taxa within this complex subgroup. Hence the subspecies included within *C. arnoldii*, that is *C. arnoldii* ssp. *arnoldii*, *C. arnoldii* ssp. *clauzadeana*, *C. arnoldii* ssp. *nana* and *C. arnoldii* ssp. *obliterata* (the latter represented by several morphotypes) were not resolved.

In the *C. saxicola* subgroup relationships among representatives of three morphospecies within clade *A* (*C. arnoldiiconfusa*, *C. rouxii* and *C. saxicola* s. str.) were resolved as three monophyletic groups with high bootstrap support only by MP5 (FIG. 4). In clade *B* each morphospecies for which more than one sequence was available (i.e. *C. biatorina*, *C. decipiens* and *C. pusilla*) was significantly recovered as monophyletic with MP5 (FIG. 4). Because the B2 dataset included several individuals of *Caloplaca schistidii* all four morphospecies of clade *B* were recovered as significantly monophyletic. However, their relationships within this clade remain uncertain.

## DISCUSSION

*Relative contribution of morphological and molecular data to the final combined analysis.*—Based on analyses of the morphological dataset, we have demonstrated that there are insufficient synapomorphies (FIGS. 2A, 4) to unravel the great majority of relationships within the *C. saxicola* group and circumscribe monophyletic taxa. However, ITS had the appropriate level of variation to resolve most relationships within this group. Phylogenetic analyses restricted to ITS data supported the monophyly of most taxa defined on the basis of traditional taxonomic practices. Yet for routine identification of taxa within this complex the most practical approach would be a diagnostic key incorporating a combination of phenotypic and ecological traits as well as geographical information, as detailed in Gaya (2009).

Morphological data (MP1) contributed to the grouping of two specimens of *C. decipiens*, two specimens of *C. arnoldii* ssp. *obliterata* and the *C. rouxii* and *C. saxicola* s. str. clade. The clade formed by *C. rouxii* and *C. saxicola* s. str. delimits two taxa characterized as having mainly short or almost no lobes, thick thalli and narrow ellipsoid spores. The close relationship of *C. biatorina* and *C. schistidii* might be due to the extremely narrow equatorial thickening of the spore shared by both species. The proximity of *C. ignea* to *C. biatorina* might be due to their strong similarity in macroscopic appearance, which probably is due to convergence. Both species have similar reddish thalli and apothecia and considerably long lobes. Whereas *C. biatorina* is typical of alpine and montane zones in Europe, North America and Asia, *C. ignea* is known only from California and it seems to be an endemic species of this maritime region (Arup 1995). Despite the poor resolving power of the morphological data when analyzed separately, significant support for *C. saxicola* s. str. was obtained only when morphological data were added to the ITS dataset. Moreover, many unequivocal morphological synapormorphies were revealed for five of the seven species of the *C. saxicola* group with more than one representative specimen when phylogenetic analyses were based on a concatenated phenotypic and ITS dataset (FIG. 4). The inclusion of phylogenies based on morphological characters that can serve as a comparison to the molecular results can provide a source of accurate independent corroboration to the trees based on molecular data and also can improve resolution (e.g. Doyle 1992, Lutzoni and Vilgalys 1995, Hillis and Wiens 2000, Jenner 2004).

*Phylogenetic recognition of treated taxa.*—The phylogeny based on concatenated datasets presented here provides support for eight of the species recognized by Gaya (2009): *Caloplaca arnoldii* s.l., *C. biatorina*, *C.*

*decipiens*, *C. arnoldiiconfusa*, *C. pusilla*, *C. rouxii*, *C. saxicola* s. str. and *C. schistidii*. The *C. arnoldii* complex includes four subspecies that are differentiated based on morphology; however their circumscription is unclear and phylogenetic relationships are not yet resolved (FIG. 4).

*Caloplaca saxicola* s. str. was difficult to delimit morphologically, and the selection of specimens representing its morphological and ecological amplitude for this study was problematic. The poor understanding of the circumscription of this species resulted in identification uncertainties, and often the herbarium material labeled as *C. saxicola* corresponded to another taxon. Unfortunately reliably identified European specimens were too old or too scarce for molecular systematic work, so we used North American material. Therefore, the phylogenetic circumscription of *C. saxicola* s. str. presented here is based mainly on North American specimens and corresponds to the morphological delimitation described by Gaya (2009). This delimitation agrees with one of the unnamed morphotypes of *C. saxicola* described by Wetmore and Kärnefelt (1998), characterized by clusters of large apothecia and extremely short lobes. In future studies sampling should be extended to other geographical regions.

*Caloplaca saxicola* has been treated commonly in a broad sense, but several subspecies also have been considered (Clauzade and Roux 1985). *C. saxicola* ssp. *pulvinata* (A. Massal.) Clauzade & Cl. Roux is one of the best known subspecies. As stated by Weddell (1876), this subspecies is so widespread that it often has been confused with typical *C. saxicola*. However, in many cases *C. saxicola* ssp. *pulvinata* has been consciously subsumed within a broader concept of the species (Poelt 1954, 1969; Wade 1965; Nordin 1972; Wetmore and Kärnefelt 1998). Indeed Clauzade and Roux (1985) are the only authors who have carried out an infraspecific treatment of *C. saxicola* in an attempt to cover the morphological variability of this taxon. *C. saxicola* ssp. *pulvinata*, for which Gaya (2009) recovered the name *C. pusilla*, is a taxon that differs from typical *C. saxicola* in many features. For example, lobes are well developed and spores are much wider in *C. pusilla* than in *C. saxicola* s. str. In this study we tried to include specimens that represented the full spectrum of morphological variability described for *C. pusilla* (Gaya 2009). These two species are well nested in the two main lineages within the *C. saxicola* subgroup—clades *A* and *B*, supporting their recognition as distinct species (FIG. 4).

Another taxon treated as a subspecies within *C. saxicola* by Clauzade and Roux (1985), *C. saxicola* ssp. *miniata*, was named *C. rouxii* by Gaya (2009). This taxon is recovered as monophyletic independently by the phylogenetic analysis of the morphological and ITS datasets when analyzed separately. The combination of both datasets further resolves the relationship of *C. rouxii* as sister of *C. saxicola* s. str., with significant statistical support (FIG. 4). The morphological features that characterize *C. rouxii* (Gaya 2009), its particular distribution and ecology and the phylogenetic results of this study (FIG. 4), all support the recognition of this taxon at the species level.

*Caloplaca arnoldiiconfusa* is another taxon with a problematic morphological circumscription, which in the works of Poelt (1954, 1969) and Poelt and Hinteregger (1993) had been mistaken and included within a broad concept of *C. arnoldii*. Clauzade and Roux (1985) recognized material referred here as *Caloplaca arnoldiiconfusa* as a subspecies of *C. saxicola* (*C. saxicola* ssp. *arnoldii*) and treated typical *C. arnoldii* as *C. saxicola* ssp. *biatorinoides*. The examination of Weddell's exsiccata by Gaya (2009) showed that *Lecanora arnoldii* corresponds to what she named *C. arnoldii* ssp. *arnoldii* (Gaya 2009), the typical species of *C. arnoldii* s.l. (corresponding to the *C. arnoldii* subgroup, FIG. 4) but not to Poelt's concept of *C. arnoldii* (Poelt 1954, 1969; see Gaya 2009). Nevertheless, Poelt's concept has been broadly used and includes specimens with well delimited rosettes and larger spores than *C. arnoldii* ssp. *arnoldii*. Consequently Gaya (2009) described it as a new species, *C. arnoldiiconfusa*. In this study *C. arnoldiiconfusa* is represented by specimens from central Europe identified by Poelt as *C. arnoldii* and also by material collected by U. Arup and used in his study of *C. demissa* (Körb.) Arup & Grube under *C. arnoldii* (Arup and Grube 1999). Based on this material, *C. arnoldiiconfusa* was recovered as a well supported monophyletic group within clade *A* of the *C. saxicola* subgroup (FIG. 4), whereas *C. arnoldii* ssp. *arnoldii* is part of the *C. arnoldii* subgroup.

Phylogenetic analyses well differentiate these two previously conflated taxa (*C. arnoldii* vs. *C. arnoldiiconfusa*). However, specimens of *C. arnoldiiconfusa* occasionally may resemble *C. arnoldii* ssp. *arnoldii* and require a closer look to detect differences in the cortex structure and spore size (see Gaya 2009).

One of the most easily recognized species of the *C. saxicola* group is *C. biatorina*. With material from Europe, North America and Asia it is one of the most widely sampled taxa in this study. *C. biatorina* in general is recognized as a species by many European lichenologists (e.g. Poelt 1954, 1969; Nordin 1972; Clauzade and Roux 1985; Nimis 1993) but was considered a morphotype of *C. saxicola* in North America (Wetmore and Kärnefelt 1998). In this study three specimens from North America, labeled *C. saxicola*, indeed belong to *C. biatorina* (specimens 7–

9, FIG. 4). In MP2, MP4, MP5 and B2 searches these three North American specimens form a well supported monophyletic group within *C. biatorina*. The sample from Pakistan conversely appears to be sister of two specimens from Spain. Of note the two latter specimens and the Asiatic sample share the same ecology; all three were collected above 2000 m, which represents the maximum elevation among the sampled specimens, except for one of the samples from North America. Alpine habitats of Europe and Asia might be where this species originated. The clear morphological distinctions and the well supported phylogenetic circumscription of *C. biatorina* make this species particularly attractive for further population and phylogeographic studies.

Of the sorediate species traditionally included within the *C. saxicola* group (*C. decipiens*, *C. cirrochroa* (Ach.) Th. Fr., *C. obliterans* (Nyl.) Blomb. & Forssell and *C. proteus* Poelt) only *C. decipiens* is truly allied with the *C. saxicola* group (Gaya et al. 2008). The two specimens of *C. decipiens* from Europe and North America are monophyletic, but its placement within clade *B* of the *C. saxicola* subgroup remains uncertain (FIGS. 3B and 4).

Finally, *C. schistidii* (previously *Fulgensia schistidii*, Kasalicky et al. 2000) falls within the *C. saxicola* group, as demonstrated by Gaya et al. (2003, 2008), more specifically within clade *B* of the *C. saxicola* subgroup. The orangish thallus, presence of lobes, septate spores as well as a cortex and medulla more compact than in *Fulgensia* species are morphological character states shared by other members of the group that support this phylogenetic placement.

Unlike the *C. saxicola* subgroup, relationships within the *C. arnoldii* subgroup (also referred here as the *C. arnoldii* species complex) remained unresolved even when combined morphological and ITS data were analyzed. None of the infraspecific taxa included in the *C. arnoldii* species complex is phylogenetically well circumscribed, nor has a geographic distribution pattern been observed, with the exception of two specimens of *C. arnoldii* ssp. *obliterata* from North America and two others from the British Isles, which appear always as monophyletic entities. Part of the material that we included here as *C. arnoldii* ssp. *obliterata* traditionally has been considered by British authors as *C. arnoldii* in a strict sense (see Gaya 2009). Two other subspecies, *C. arnoldii* ssp. *clauzadeana* and *C. arnoldii* ssp. *nana*, are also part of the *C. arnoldii* subgroup. Because each is represented by a single sequence (from the type material) we cannot draw further conclusions about the monophyly of these two taxa.

In the *C. arnoldii* subgroup historic taxonomic and nomenclatural confusion and the difficulty of delimiting its taxa phenotypically is matched by a lack of phylogenetic resolution in our analyses. Most of the branches supporting relationships among taxa within the *C. arnoldii* subgroup are similar or shorter than those observed among specimens within a single species of the *C. saxicola* subgroup (FIG. 3A, B). This could indicate a faster rate of speciation, which could not be detected by ITS, or an evolutionary process where extensive gene flow is persisting, preventing genetic isolation and therefore speciation.

In some cases the subspecies included within the *C. arnoldii* subgroup are sympatric; for example, *C. arnoldii* ssp. *arnoldii*, *C. arnoldii* ssp. *clauzadeana* and *C. arnoldii* ssp. *nana* all grow on vertical walls and overhangs protected from direct rainwater runoff, usually avoiding direct sunlight, and mainly on calcareous substrata. *C. arnoldii* ssp. *arnoldii* and *C. arnoldii* ssp. *nana* are distributed across Europe and can overlap with *C. arnoldii* ssp. *clauzadeana* in the south of France. *Caloplaca arnoldii* ssp. *obliterata* has a more northern distribution (abundant in Scandinavia and British Isles), but some of its morphotypes can overlap with the areas of distribution of the previous subspecies. This is the only subspecies with a different ecology because it usually grows on siliceous substrata. Furthermore, when these subspecies grow together, individuals with intermediate character states, especially regarding thallus color, can appear occasionally. Further taxonomic studies on the *C. arnoldii* subgroup would require multiple faster evolving molecular markers and an extensive sampling at the population level.

*More characters but less efficient phylogenetic methods versus fewer characters but more efficient phylogenetic methods.*—In general the topologies inferred by the methodologies (MP, ML and Bayesian) from the ITS data differed slightly regarding the delimitation of most taxa and the relationships among them. In all ITS analyses and in the combined search the two main lineages (*C. arnoldii* and *C. saxicola* subgroups) were recovered. Therefore, ITS nrDNA, with the inclusion of INAASE characters and morphology, phylogenetically delimited the morphospecies evaluated and previously treated by Gaya (2009). In some cases the signal from ambiguously aligned regions recovered with character coding (INAASE) or a Bayesian approach (BAli-Phy) permitted the recovery of population structure information, as in *C. biatorina* from North America.

Working with such closely related species suggested that the inclusion of INAASE characters would contribute less to phylogenetic resolution and support, compared to phylogenetic analyses conducted at the family level by Gaya et al. (2003, 2008), because the percentage of ambiguously aligned sites was lesser.

Nevertheless, we observed again that the inclusion of these characters in three of the four parsimony analyses increased resolution and support within the *C. saxicola* group (TABLE II). However the Bayesian approach implemented in BAli-Phy, which allows taking into account alignment uncertainty by integrating over all alignments in proportion to their posterior probabilities, has proved to perform nearly as well as MP analyses when recovering phylogenetic signal from ambiguously aligned regions with INAASE and adding more characters (morphology). This Bayesian analysis of nonaligned sequences (BAli-Phy), which captures all available signal from the two transcribed spacers (contrary to character coding methods), contributed more to the degree of support than any other phylogenetic method implemented in this study except MP5, which also included morphological characters. Of note this is the first study where this type of Bayesian analysis has been implemented on a dataset of more than 50 sequences, which previously had been limited by compute time demands. For this study we constrained part of the sequences (the conserved 5.8S) as aligned and fixed, contrary to the ITS1 and ITS2, which were not aligned a priori for the B2 analysis. This approach considerably reduced the compute time compared to feeding completely unaligned sequences into the Bayesian analysis performed with BAli-Phy. The analyses took 31–34 d and used six computers, one for each MCMC run.

Maximum likelihood (ML) and traditional Bayesian analyses as implemented in MrBayes conversely did not show greater resolution than the MP analyses or the approach described above when recovering signal from ambiguously aligned regions. The number of significantly supported internodes was fewer (15 for ML and 22 for Bayesian vs. 29 for weighted MP with INAASE characters and 32 for B2 with alignment estimation, TABLE II). Therefore, the availability of the program BAli-Phy offers an alternative to the previous dilemma of choosing between more characters but less efficient phylogenetic methods (e.g. character coding of ambiguously aligned regions analyzed with MP) versus fewer characters but more efficient phylogenetic methods (i.e. exclusion of ambiguously aligned regions and use of ML and Bayesian methods) by offering the possibility of including all sites of ambiguously aligned regions and using a Bayesian method. One remaining technical limitation of BAli-Phy is the inability to analyze morphological and molecular data simultaneously.

## ACKNOWLEDGMENTS

The authors thank the curators of herbaria BCN, E, GZU, LD, MARSSJ, MIN, VAB-LICH, in addition to Ulf Arup and Claude Roux, for providing material used in this study. Claude Roux is warmly thanked for sharing with us his knowledge about the *Caloplaca saxicola* group. We also thank Ulf Arup and Ulrik Søchting for helpful comments; Frank Kauff and Stefan Zoller for writing computer programs useful for this study; Carl Rothfels for the English revision and proofreading; and Nestor Hladun for providing lab facilities. This work was carried out within projects (PB 96-1115-C04-02, BOS 2001-0869-C04-02, CGL 2005-04322), Spain, and CIRIT (2001 SGR 00095, 2005 SGR 01047), Catalonia. We gratefully acknowledge support from a doctoral scholarship (FI-programa propi, Universitat de Barcelona), a postdoctoral grant from the Fulbright Scholar Program and Spanish government to EG and a grant from the National Science Foundation (DEB-0919455) to FL. BDR was supported by the National Evolutionary Synthesis Center (NSF EF-0905606).

## LITERATURE CITED

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. Proceedings of the 2nd international symposium on information theory. Budapest: Akademiai Kiado. p 267–281.

Arup U. 1995. Eight species of *Caloplaca* in coastal western North America. Bryologist 98:92–111, doi:10.2307/3243645

———. 2006. A new taxonomy of the *Caloplaca citrina* group in the Nordic countries, except Iceland. Lichenologist 38:1–20, doi:10.1017/S0024282905005402

———. 2009. The *Caloplaca holocarpa* group in the Nordic countries, except Iceland. Lichenologist 41:111–130, doi:10.1017/S0024282909008135

———, Grube M. 1999. Where does *Lecanora demissa* (Ascomycota, Lecanorales) belong? Lichenologist 31: 419–430.

Beiko RG, Keith JM, Harlow TJ, Ragan MA. 2006. Searching for convergence in phylogenetic Markov chain Monte Carlo. Syst Biol 55:553–565, doi:10.1080/10635150600812544

Calinski T, Harabasz J. 1974. A dendrite method for cluster analysis. Commun Stat 3:1–27, doi:10.1080/03610927408827101

Cao Y, Adachi J, Janke A, Pääbo S, Hasegawa M. 1994. Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. J Mol Evol 39:519–527, doi:10.1007/BF00173421

Cartwright RA. 2009. Problems and solutions for estimating indel rates and length distributions. Mol Biol Evol 26: 473–480, doi:10.1093/molbev/msn275

Clauzade G, Roux C. 1985. Likenoj de Okcidenta Europo. B Soc Bot Centre-Ouest, Numéro Special 1:1–893.

De Cáceres M, Font X, García R, Oliva F. 2003. VEGANA, un paquete de programas para la gestión y análisis de datos ecológicos. VII Congreso Nacional de la Asociación Española de Ecología Terrestre. Barcelona, July 2003:1484–1497.

———, Oliva F, Font X, Vives S. 2007. GINKGO, a program

for non-standard multivariate fuzzy analysis. Adv Fuzzy Set Syst 2:41–56.

Doyle JJ. 1992. Gene trees and species trees—molecular systematics as one-character taxonomy. Syst Bot 17:144–163, doi:10.2307/2419070

Felsenstein J. 1981. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. Biol J Linn Soc 16:183–196, doi:10.1111/j.1095-8312.1981.tb01847.x

———. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791, doi:10.2307/2408678

Gardes M, Bruns TD. 1993. ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. Mol Ecol 2:113–118, doi:10.1111/j.1365-294X.1993.tb00005.x

Gaya E. 2009. Taxonomical revision on the *Caloplaca saxicola* group (Teloschistaceae, lichen-forming Ascomycota). Bibl Lichen 101. Berlin, Suttgart: J. Cramer.

———, Lutzoni F, Zoller S, Navarro-Rosinés P. 2003. Phylogenetic study of *Fulgensia* and allied *Caloplaca* and *Xanthoria* species (Teloschistaceae, lichen-forming ascomycota). Am J Bot 90:1095–1103, doi:10.3732/ajb.90.7.1095

———, Navarro-Rosinés P, Llimona X, Hladun N, Lutzoni F. 2008. Phylogenetic reassessment of the Teloschistaceae (lichen-forming Ascomycota, Lecanoromycetes). Mycol Res 112:528–546, doi:10.1016/j.mycres.2007.11.005

Goldman N, Whelan S. 2002. A novel use of equilibrium frequencies in models of sequence evolution. Mol Biol Evol 19:1821–1831.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174, doi:10.1007/BF02101694

Hillis DM, Heath TA, John KS. 2005. Analysis and visualization of tree space. Syst Biol 54:471–482, doi:10.1080/10635150590946961

———, Wiens JJ. 2000. Molecules versus morphology in systematics: conflicts, artifacts and misconceptions. In: Wiens JJ, ed. Phylogenetic analysis of morphological data. Washington, DC: Smithsonian Institution Press. p 1–19.

Huelsenbeck JP, Crandall KA. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu Rev Ecol Syst 28:437–466, doi:10.1146/annurev.ecolsys.28.1.437

———, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinf Appl Note 17:754–755.

Jenner RA. 2004. Accepting partnership by submission? Morphological phylogenetics in a molecular millennium. Syst Biol 53:333–342, doi:10.1080/10635150490423962

Kärnefelt I. 1989. Morphology and phylogeny in the *Teloschistales*. Cryptogam Bot 1:147–203.

Kasalicky T, Döring H, Rambold G, Wedin M. 2000. A comparison of ITS and LSU nrDNA phylogenies of *Fulgensia* (Teloschistaceae, Lecanorales), a genus of lichenized ascomycetes. Can J Bot 78:1580–1589, doi:10.1139/cjb-78-12-1580

Kirk PM, Cannon PF, Minter DW, Stalpers JA. 2008. Dictionary of the Fungi. 10th ed. Wallingford, UK: CAB International. 784 p.

Lanave C, Preparata G, Saccone C, Serio G. 1984. A new method for calculating evolutionary substitution rates. J Mol Evol 20:86–93, doi:10.1007/BF02101990

Lohtander K, Källersjö M, Moberg R, Tehler A. 2000. The family *Physciaceae* in Fennoscandia: phylogeny inferred from ITS sequences. Mycologia 92:728–735, doi:10.2307/3761429

Lunter G. 2007. Probabilistic whole-genome alignments reveal high indel rates in the human and mouse genomes. Bioinformatics 23:i289–i296, doi:10.1093/bioinformatics/btm185

Lutzoni F, Vilgalys R. 1995. Integration of morphological and molecular datasets in estimating fungal phylogenies. Can J Bot 73:S649–S659, doi:10.1139/b95-307

———, Wagner P, Reeb V, Zoller S. 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst Biol 49:628–651, doi:10.1080/106351500750049743

MacQueen J. 1967. Some methods for classification and analysis of multivariate observation. In: Le Cam LM, Neyman J, eds. Proceedings of the 5th Berkeley symposium on mathematical statistics and probability. Berkeley: Univ. California Press. p 281–297.

Maddison WP, Maddison DR. 2001. MacClade manual: analysis of phylogeny and character evolution. 4.01. Sunderland, Massachusetts: Sinauer Associates.

Mason-Gamer RJ, Kellogg EA. 1996. Testing for phylogenetic conflict among molecular dataset in the tribe *Triticiae* (Gramineae). Syst Biol 45:524–545, doi:10.1093/sysbio/45.4.524

Muggia L, Grube M, Tretiach M. 2008. A combined molecular and morphological approach to species delimitation in blackfruited, endolithic *Caloplaca*: high genetic and low morphological diversity. Mycol Res 112:36–49, doi:10.1016/j.mycres.2007.02.001

Nimis PL. 1993. The Lichens of Italy. An annotated catalogue. Monografia XII. Torino: Mus Reg Sci Nat.

Nordin I. 1972. Caloplaca sect. Gasparrinia i Nordeuropa. Taxonomiska och Ekologiska Studier [doctoral dissertation]. Uppsala, Sweden: Skriv Service AB.

Poelt J. 1954. Die Gelappten arten der flechtengattung *Caloplaca* in Europa. Mitt Bot Staatssamml München 11:11–31.

———. 1969. Bestimmungsschlüssel europäischer Flechten. Lehre, Vaduz: J. Cramer.

———, Hinteregger E. 1993. Beiträge zur Kenntnis der Flechtenflora des Himalaya VII. Die Gattungen Caloplaca, Fulgensia und Ioplaca (mit englischem Bestimmungsschlüssel). Bibl Lichen 50. Berlin, Stuttgart: J. Cramer.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinf Appl Note 14:817–818.

Prendini L. 2001. Species or supraspecific taxa as terminals in cladistic analysis? Ground plans versus exemplars revisited. Syst Biol 50:290–300, doi:10.1080/10635150118650

Redelings BD, Suchard MA. 2005. Joint Bayesian estimation of alignment and phylogeny. Syst Biol 54:401–418, doi:10.1080/10635150590947041

———, ———. 2007. Incorporating indel information into phylogeny estimation for rapidly emerging pathogens. BMC Evol Biol 7:40, doi:10.1186/1471-2148-7-40

Robinson DF, Foulds LR. 1981. Comparison of phylogenetic trees. Math Biosci 53:131–147, doi:10.1016/0025-5564(81)90043-2

Søchting U, Lutzoni F. 2003. Molecular phylogenetic study at the generic boundary between the lichen-forming fungi *Caloplaca* and *Xanthoria* (Ascomycota, Teloschistaceae). Mycol Res 107:1266–1276, doi:10.1017/S0953756203008529

Suchard M, Kitchen C, Sinsheimer J, Weiss R. 2003. Hierarchical phylogenetic models for analyzing multipartite sequence data. Syst Biol 52:649–664, doi:10.1080/10635150390238879

———, Redelings BD. 2006. BAli-Phy: simultaneous Bayesian inference of alignment and phylogeny. Bioinformatics 22:2047–2048, doi:10.1093/bioinformatics/btl175

Swofford DL. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526.

Verseghy K. 1970. Hazai *Gasparrinia* fajok I. Bot Közlem 57: 23–29.

———. 1971. Hazai *Gasparrinia* fajok II. Rendszertani resz. Bot Közlem 58:21–28.

———. 1972. Hazai *Gasparrinia* fajok III. Rendszertani rez (befejezes). Bot. Közlem. 59:13–18.

Vilgalys R, Hester M. 1990. Rapid identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J Bacteriol 172:4238–4246.

Vondrák J, Říha P, Arup U, Søchting U. 2009. The taxonomy of the *Caloplaca citrina* group (Teloschistaceae) in the Black Sea region; with contributions to the cryptic species concept in lichenology. Lichenologist 41:571–604, doi:10.1017/S0024282909008317

———, Šoun J, Hrouzek P, Říha P, Kubásek J, Palice Z, Søchting U. 2008. *Caloplaca subalpina* and *C. thracopontica*, two new saxicolous species from the *Caloplaca cerina* group (Teloschistaceae). Lichenologist 40:375–386.

Wade AE. 1965. The genus *Caloplaca* Th. Fr. in the British Isles. Lichenologist 3:1–28, doi:10.1017/S0024282965000038

Weddell HA. 1876. Notice monographique sur les *Amphiloma* de la flore française. B Soc Bot Fr 23(CR des Séanc.): 83–99.

Wetmore CM, Kärnefelt I. 1998. The lobate and subfruticose species of *Caloplaca* in North and Central America. Bryologist 101:230–255.

Wheeler WC. 1990. Combinatorial weights in phylogenetic analysis: a statistical parsimony procedure. Cladistics 6: 269–275, doi:10.1111/j.1096-0031.1990.tb00543.x

White TJ, Bruns TD, Lee SB, Taylor JW. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. PCR protocols: a guide to methods and applications. New York: Academic Press. p 315–322.

Zwickl DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion [doctoral dissertation]. Austin: Univ. Texas Press. 115 p.