E. SYSTEMATICS / SYSTÉMATIQUE

E1.1.

Integration of morphological and molecular data sets in estimating fungal phylogenies

François Lutzoni and Rytas Vilgalys

Abstract: To provide a clearer picture of fungal species relationships, increased efforts are being made to include both molecular and morphological data sets in phylogenetic studies. This general practice in systematics has raised many unresolved questions and controversies regarding how to best integrate the phylogenetic information revealed by morphological and molecular characters. This is because phylogenetic trees derived using different data sets are rarely identical. Such discrepancies can be due to sampling error, to the use of an inappropriate evolutionary model for a given data set, or to different phylogenetic histories between the organisms and the molecule. Methods have been developed recently to test for heterogeneity among data sets, although none of these methods have been subjected to simulation studies. In this paper we compare three tests: a protocol described by Rodrigo et al., an adapted version of Faith's T-PTP test, and Kishino and Hasegawa's likelihood test. These tests were empirically compared using seven lichenized and nonlichenized *Omphalina* species and the related species *Arrhenia lobata* (Basidiomycota, Agaricales) for which nrDNA large subunit sequences and morphological data were gathered. The results of these three tests were inconsistent, Rodrigo's test being the only one suggesting that the two data sets could be combined. One of the three most parsimonious trees obtained from the combined data set with eight species is totally congruent with the relationships among the same eight species in an analysis restricted to the same portion of the nrDNA large subunit but extended to 26 species of *Omphalina* and related genera. Therefore, the results from phylogenetic analyses of this large molecular data set converged on one of the three most parsimonious topologies generated by the combined data set analysis. This topology was not recovered from either data set when analysed separately. This suggests that Rodrigo's homogeneity test might be better suited than the two other tests for determining if trees obtained from different data sets are sampling statistics of the same phylogenetic history.

Key words: data sets heterogeneity, homogeneity test, lichen phylogeny, *Omphalina*, ribosomal DNA.

Résumé : Afin de constituer une meilleure représentation des relations entre les espèces fongiques, de plus en plus d'efforts sont déployés pour inclure à la fois les caractères moléculaires et les caractères morphologiques dans les études phylogénétiques. Cette pratique générale en systématique a soulevé beaucoup de questions nouvelles et de controverses sur les meilleures façons d'intégrer les informations phylogénétiques provenant des caractères morphologiques et moléculaires. Ceci provient du fait que les dendrogrammes phylogénétiques dérivés des deux ensembles de données sont rarement identiques. De telles différences peuvent provenir d'erreurs d'échantillonage, de l'utilisation de modèles évolutionnaires inappropriés pour les données en cause, ou à des histoires phylogénétiques différentes entre les organismes et leurs molécules. Des méthodes ont été récemment développées pour vérifier l'hétérogénéité d'ensembles de données, bien qu'aucune de ces méthodes n'ait été soumise à des études de simulation. Dans leur présentation, les auteurs comparent trois méthodes d'évaluation : un protocole décrit par Rodrigo et al., une version adaptée du test T-PTP de Faith et le test de vraisemblance de Kishino et Hasegawa. Ces tests ont été empiriquement comparés en utilisant sept espèces d'*Omphalina* lichénisées ou non-lichénisées et l'espèce voisine *Arrhenia lobata* (Basidiomycota, Agaricales) pour lesquelles les séquences de la grande sous-unité du nrADN et les données morphologiques ont été obtenues. Les résultats de ces trois tests ne concordent pas, celui de Rodrigo étant le seul à suggérer que les deux ensembles de données pourraient être combinés. Un des trois arbres les plus parcimonieux obtenus à partir d'une combinaison de l'ensemble des données incluant huit espèces est complètement congruent avec les relations entre les

Received August 19, 1994.

F. Lutzoni[1] **and R. Vilgalys.** Department of Botany, Duke University, Durham, NC 27708, U.S.A.

[1] Author to whom all correspondence should be addressed.

mêmes huit espèces dans une analyse restreinte à la même partie de la grande sous-unité, mais étendue à 26 espèces d'*Omphalina* et de genres apparentés. Conséquement les résultats des analyses phylogénétiques de ce grand ensemble de données moléculaires convergent vers une des trois représentations topologiques les plus parcimonieuses générées par l'analyse des données prises dans leur ensemble. Cette topologie n'a pas pu être révélée lorsque les données ont été analysées séparément. Ceci sugère que le test d'homogénéité de Rodrigo pourrait être plus approprié que les deux autres tests pour déterminer si les dendrogrammes obtenus à partir de différentes bases de données constituent des échantillonnages statistiques de la même histoire phylogénétique.

Mots clés : hétérogénéïté des ensembles de données, tests d'homogénéïté, phylogénie des lichens, *Omphalina*, ADN ribosomal.
[Traduit par la rédaction]

Introduction

The most fundamental goal of systematists working on fungi is to infer the evolutionary history that led to the extant species of fungi. The first attempts were done using morphological characters, but several limitations were encountered. One major obstacle was the difficulty of implementing the recognition criteria of homology, i.e., position, quality of resemblance, and continuance of similarity through intermediate species (Wiley 1981) when working with morphology at high taxonomic levels. Another significant impediment was the lack of morphological characters for microscopic fungi. An additional complication was imposed by species having an anamorphic stage. Chemotaxonomic studies on fungi were burdened with the same problems of determining homology and lack of characters, with additional problems related to the association of secondary compounds with biosynthetic pathways (Gowan 1989). With the exception of Tehler's work (1988), these constraints restricted the contribution of morphological and chemotaxonomic studies to phylogenetic questions below the ordinal level and to macrofungi.

With the development of molecular biology and the transfer of molecular techniques to systematics, the homology problem seemed less stringent. The lack of molecular characters can be solved by discovering new regions of genomes that are appropriate to addressing specific phylogenetic questions. The major contribution of molecular systematics to mycology was to offer the opportunity to formulate, often for the first time, phylogenetic hypotheses for microfungi (Kurtzman and Robnett 1991; Fell et al. 1992; Nishida and Sugiyama 1993), for the major groups of fungi (Berbee and Taylor 1992*a*; Bruns et al. 1992; Spatafora and Blackwell 1993*a*, 1993*b*), and for anamorphs and their respective teleomorphs (Guadet et al. 1989; Berbee and Taylor 1992*b*; Rehner and Samuels 1994; LoBuglio and Taylor 1993). However, molecular systematics will not solve all problems associated with estimating organismal phylogenies. Despite the homology problem being less prominent with molecular data, especially in the case of nucleotide and amino acid sequences where the concept of positional homology applies, there is still room for mistaken inferences about homologous molecular relationships (Hillis 1994). Even if the potential to find new informative molecular characters is tremendous, a given molecular data set may lack characters necessary to solve important parts of the phylogenetic history of a given group of fungi.

Another limitation of molecular data to recover species/population phylogenetic history is the potential for differential evolution between protein or nucleotide sequences and organisms (Hillis 1987; Nei 1987; Doyle 1992). The phylogenetic history for a given molecule can be different from the organismal phylogeny owing to lineage sorting acting on polymorphic molecular characters, lateral transfer of parts of the genome, and hybridization. Since fungal systematists are primarily interested in species/population phylogenies rather than gene phylogenies, this problem of differential phylogenetic history between molecular characters and morphological characters justifies the importance of using both types of data to estimate organismal phylogenies. In addition to gaining a better idea of the robustness of a given phylogenetic hypothesis, congruence between morphological and molecular data sets becomes a criterion for determining if a particular set of molecular characters can be used to estimate a species/population phylogeny for a given set of taxa.

Another reason why molecular and morphological trees obtained from the same group of taxa may differ is the failure of a phylogenetic method to recover the right topology for one or both data sets (Bull et al. 1993). There are several factors that can cause a phylogenetic method to fail to recover the right phylogenetic tree. For instance, the molecular data set can be saturated by change and, therefore, not contain sufficient signal necessary to recover the phylogenetic history (Archie 1989; Faith and Cranston 1991; Hillis 1991; Hillis and Huelsenbeck 1992; Alroy 1994). Another problem might be sampling error, both for taxa selection and amount of data. However, even if we had unlimited data, there is at present no algorithm, nor computer powerful enough, to guarantee recovery of the most parsimonious or most likely phylogenetic estimate from very large data sets (Hendy and Penny 1982). Trees with long terminal branches and short internodes can cause phylogenetic methods to be inconsistent, i.e., as more data are gathered the methods converge on the wrong topology (Felsenstein 1978). Finally, the assumptions underlying the evolutionary model implemented when using a phylogenetic method might not fit the data (Bull et al. 1993; Goldman 1993; Rodrigo et al. 1993).

Despite the need to integrate morphological and molecular data sets in estimating species/population phylogenies, this approach has raised many new controversies and unresolved questions (Miyamoto 1985; Kluge 1989; Swofford 1991; Bull et al. 1993; de Queiroz 1993; Eernisse and Kluge 1993; Rodrigo et al. 1993; Chippindale and Wiens 1994; Huelsenbeck et al. 1994). This is mainly because phylogenetic trees derived using different data sets are rarely identical, and as systematists we are interested in molecular data only if the molecular trees are sampling statistics of the

organismal phylogeny. The problem becomes even more acute when the conflicting trees based on morphological and molecular data are strongly supported by their respective data set. There are two major schools of thought on how to address this problem (see Bull et al. 1993 for references to different approaches). One suggests that the molecular and morphological data sets collected from the same taxa should be combined prior to any phylogenetic analysis (total evidence argument) while the other argues that the data should be analyzed separately first. The proponents of the second approach can be separated into three subgroups, where conflicting topologies should (*i*) never be combined, (*ii*) be combined using consensus methods, and (*iii*) be tested first (homogeneity test) to determine if the incongruence between the two data sets is due to sampling error (Bull et al. 1993; Rodrigo et al. 1993).

It has been shown that a phylogenetic estimation is not necessarily improved by combining data sets (Bull et al. 1993) and a consensus tree obtained from trees based on different data sets is not necessarily the most conservative estimate of a species/population phylogeny (Barrett et al. 1991). Therefore, data sets need to be analyzed separately and tested to determine if the conflict between data sets, if there is conflict, is due mainly to sampling error (Bull et al. 1993; Rodrigo et al. 1993). If the null hypothesis cannot be rejected, i.e., if the molecular and morphological trees are no more different than would be expected owing to sampling error, then the data sets can be combined. This decision to combine data sets is in agreement with the proponents of the total evidence argument, but the rationale is different (see Miyamoto 1985; Kluge 1989).

However, when the null hypothesis is rejected, i.e., when different data sets yield significantly different phylogenetic estimates that are not due to sampling error, it is better to extend the analysis on separate data sets before making any decision about combining them. If the difference between the morphological and molecular trees is due to inappropriate assumptions by the reconstruction model for one or both data sets (Goldman 1993), a different reconstruction model should be used with a better fit to the data and the test repeated (Bull et al. 1993). If the difference between the species/population tree and the gene tree is real (i.e., the molecular and morphological data have a different phylogenetic history), the phylogeny for the entire set of taxa cannot be resolved. However, it is possible that parts of the morphological and molecular trees share a common history. To find parts of the data sets that are homogeneous, Rodrigo et al. (1993) suggest identifying taxa with potentially confounding effects on parsimony and to prune these taxa one by one and repeat the test until the null hypothesis is accepted. The remaining parts of each data set can then be combined.

Rodrigo et al. (1993) developed a protocol using different tests to address specific questions related to heterogeneity of data sets. Bull et al. (1993) pointed out that Faith's (1991) T-PTP test (topology-dependent cladistic permutation tail probability test) shows potential for testing for homogeneity between data sets. Kishino and Hasegawa's (1989) likelihood test, as implemented in DNAML version 3.53c (PHYLIP, Felsenstein 1993), can be used to determine if the best phylogenetic estimate based on morphological data is not significantly worse than the best molecular tree to explain the sampled nucleotide sequences. None of these methods have been subjected to simulation studies, and therefore, we do not know how biased these tests are. In this paper we are making an empirical comparison of Rodrigo et al. (1993) method, Faith's (1991) T-PTP test, and Kishino and Hasegawa's (1989) likelihood test using lichenized and nonlichenized *Omphalina* species and the related species *Arrhenia lobata* (Basidiomycota, Agaricales) for which both molecular and morphological data were gathered.

Materials and methods

The collection data, the choice of taxa, and the choice of characters for the populations of *Omphalina* species and *Arrhenia lobata* that were used to build the morphological data set used in the present study (Tables 1 and 2) are described in Lutzoni and Vilgalys (1995). The phylogenetic analysis of the morphological data is essentially the same as described in that paper, with the exception of using version 3.1.1 of PAUP (Swofford 1993). *Omphalina rivulicola* was chosen as the outgroup for all analyses based on a broader study (F. Lutzoni and R. Vilgalys, in preparation) including 26 species of *Omphalina* and related genera using ≈ 1.1 kb from the 5′ end of the nuclear encoded large subunit ribosomal DNA.

The nucleotide sequences (≈ 1.4 kb) from the 5′ end of the nuclear encoded large subunit rDNA were obtained as described in Lutzoni and Vilgalys (1995; in preparation). The sequence fragments for each species were assembled using Sequencher version 2.1 (Gene Codes Corporation). The sequences were aligned manually using the same program and the alignment was exported as a Nexus file. The parsimony analyses on the nucleotide sequences were done using PAUP version 3.1.1 (Swofford 1993). The regions of the alignment that were ambiguous because of the presence of indels were excluded (Fig. 1). The remaining indels were treated as a fifth character state. An exhaustive search was performed on all unweighted informative sites. The robustness of the internodes was assessed by 1000 bootstrap replications (Felsenstein 1985) and decay analysis (Mishler et al. 1991). The neighbor-joining and maximum likelihood phylogenetic analyses were performed using DNADIST–NEIGHBOR and DNAML programs, respectively (PHYLIP version 3.53c; Felsenstein 1993) on the entire sequence data but excluding all positions with indels. For comparison, the neighbor-joining algorithm was implemented using Jukes and Cantor's (1969), Kimura's (1980), and Kishino and Hasegawa's (1989) models of nucleotide substitution with a transition/transversion ratio of 2.0, one category of substitution rates, and randomization of the input order of sequences. To estimate how well the internodes were supported by the molecular data, the SEQBOOT(1000 replicates)–DNADIST–NEIGHBOR programs (PHYLIP) were implemented. The majority rule consensus of these 1000 bootstrap trees was obtained using PAUP. The DNAML program was implemented using the Kishino and Hasegawa (1989) model, with the same settings as for the neighbor-joining analyses but using the global rearrangements algorithm. This was run 10 times with a different order of entry of the sequences in a stepwise addition process.

Rodrigo's protocol was implemented as described in Rodrigo et al. (1993) with the exception of using PAUP

Table 1. Characters used for parsimony analyses of seven species of *Omphalina* and *Arrhenia lobata* (from Lutzoni and Vilgalys 1995, reproduced with permission of Cryptogam. Bot., Vol. 5, pp. 71–81, © 1995 Gustav Fischer Verlag).

Character No.	Character	Character states
1	Lichenization*	0 = absent 1 = globular (*Botrydina* type) 2 = squamulose (*Coriscium* type)
2	Clamps at the base of basidium	0 = absent 1 = present
3	Proportion of basidia with different numbers of spores	0 = 70% tetrasporic, 7.5% trisporic, 15% bisporic, 7.5% unisporic 1 = 70% tetrasporic, 15% trisporic, 15% bisporic 2 = 30% tetrasporic, 70% bisporic 3 = 100% tetrasporic
4	Reproduction*	0 = potentially parthenogenic 1 = sexual
5	Pileus surface	0 = free ends not forming scales 1 = free ends forming scales
6	Cutis texture (based on Korf's (1958) classification of fungal tissue texture)	0 = epidermoidea 1 = intricata 2 = porrecta 3 = prismatica
7	Cutis-free terminal cell	0 = absent 1 = present
8	Pileus micro-incrustation on the cell wall	0 = absent 1 = present
9	Pileus laciniate incrustation on the cell wall	0 = absent 1 = present
10	Axenic culture on MEA*	0 = not possible 1 = possible
11	Pileus colour	0 = reddish-brown, becoming dark brown, becoming yellowish brown, or ± red brown to grey brown to beige 1 = pale yellow to brilliant orange yellow, or bright yellow, or olive brown when young becoming yellowish brown to yellow 2 = dark grey brown to dark purplish brown, or smoky grey brown
12	Lamellae colour	0 = pale brown 1 = bright yellow to bright orange yellow 2 = dark greyish brown to brownish grey, or whitish to pale beige, or pale grey brown, or brownish, whitish, cream, yellowish, or pale greyish 3 = grey-white to grey
13	Stipe base pubescence	0 = absent 1 = present
14	Stipe interior	0 = hollow 1 = solid
15	Stipe colour	0 = pale reddish-brown becoming medium brown, or ± red brown to grey brown to beige 1 = white to pale orange yellow, sometimes with violet tinge, or bright yellow, or olive brown at apex when young becoming yellowish to yellow or only fading 2 = dark grey brown to dark purplish brown, or smoky grey brown
16	Proportion of basidiospores with different numbers of nuclei	0 = 13% uninucleate, 60% binucleate, 13% trinucleate, and 13% > trinucleate 1 = 100% binucleate 2 = 100% uninucleate
17	Stipe	0 = absent 1 = present

NOTE: All characters were unweighted and unordered.

*These characters associated with lichenization were not used in reconstructing the phylogeny, so as to not bias the analysis toward the monophyly of lichenized species.

Table 2. Morphological–anatomical data matrix for seven species of *Omphalina* and *Arrhenia lobata*.

Species	Character No.																
	1*	2	3	4*	5	6	7	8	9	10*	11	12	13	14	15	16	17
Arrhenia lobata	0	(01)	3	1	1	(12)	0	1	1	1	2	2	?	?	?	2	0
O. epichysium	0	1	?	1	(01)	2	1	(01)	0	1	2	3	?	?	2	2	1
O. ericetorum	1	0	0	0	0	(12)	0	0	0	0	1	2	1	(01)	1	0	1
O. hudsoniana	2	0	1	0	0	(01)	1	0	0	0	1	1	1	(01)	1	2	1
O. luteovitellina	1	0	1	0	0	1	0	0	0	0	1	1	0	1	1	1	1
O. obscurata	0	1	3	1	1	(12)	1	1	1	1	2	2	0	1	2	0	1
O. rivulicola	0	1	3	1	0	(123)	0	1	1	1	0	0	(01)	0	0	2	1
O. velutina	1	0	2	0	(01)	1	0	(01)	0	0	0	2	1	1	0	?	1

NOTE: Unknown or nonapplicable character states for taxa are indicated by a question mark. Parentheses are used to accommodate taxa that were polymorphic for a given character (Maddison and Maddison 1992; Swofford 1993).

*Characters associated with lichenization excluded from the phylogenetic analysis, so as to not bias the analysis toward the monophyly of lichenized species.

Fig. 1. Nucleotide sites (62) used in this study obtained from the ≈1.4 kb sequences at the 5′ end of the nrDNA large subunit. Uninformative sites and ambiguous regions of the alignment due to the presence of indels were excluded. The numbers represent the positions of nucleotides in the original alignment.

```
                                                                  1111111111111
                  1111111222223333444444445555555566667777777788800001111233333
                  4800014480115879993444466702347778237901234699004888900096366 7
Taxon             1046742853358193789124529096330153437592416517257467156788259 6
-------------------------------------------------------------------------------
Omphalina epichysium ATT--TCGAA?CTCGTCCCTGTTAT-GTTGG---TTC-TC-GGG-GGTC????????????
O. ericetorum        ACT-CCCGTGAACCGTCGCCTTCACC-CCGG--GTTTATT--GGCTATTACCAGCTCTTGTC
O. hudsoniana        ATCT-CAAGAGTTGCTTTCTGCAGTC-TCGTTTGTCCTCTG-GGCAGCTACCTGCTCTTCTC
O. luteovitellina    GCCTCCAACAGTTGCTTTCCGCCACCCTTGTTTGTCCATTCGAGCGGTTACGAGCTCCCCCC
O. obscurata         GCTT-TTGAAACTCGCTTCCATCACTCTTAG--GTTCAAC--AATAGTCGGGTTTCCTTCTT
O. rivulicola        ACTT-CTGAAAATCGTTCTTGTTACC-CTGG---ATTTCT--GGCAATTACCTGCTTTTGTC
O. velutina          ATCTCCAACGGCTGCCCCTCGTGACC-CTGG--GAGCACTG-GGTGGCTACCTGCTCCCGCC
Arrhenia lobata      ATTT-CTGGAACCCGTCTCCATTGCTCTTAG--GTTCATC--GATTGTCGGGTTTCTTTGTT
```

version 3.1.1 (Swofford 1993) rather than COMPONENT version 1.5 (Page 1989). Another variation from Rodrigo's (1993) protocol was to combine the data sets rather than use Nelson's consensus method (sensu Page 1989) when the separate tree estimates based on morphological and molecular data were found to be topological statistics of the same phylogenetic history. Faith's (1991) a priori T-PTP test was adapted following Bull et al. (1993) guidelines. The data sets were permuted 100 times using the permute option of the SEQBOOT program of PHYLIP version 3.53c (Felsenstein 1993). The tree lengths for the most parsimonious topology based on morphology and the one based on nucleotide sequences were calculated for each permuted data set using PAUP in a batch mode (version 3.1.1, Swofford 1993). This procedure was done on the 100 permuted morphological data sets and on the 100 permuted molecular data sets. The likelihood test was performed using DNAML (PHYLIP version 3.53c) as described by Felsenstein (1993). Rodrigo's test and the adapted version of the T-PTP test were performed only on the nucleotide sites present in Fig. 1, and on the 13 informative and nonexcluded morphological characters (Table 2). The likelihood test was applied to the complete sequence data, but excluding all positions with indels.

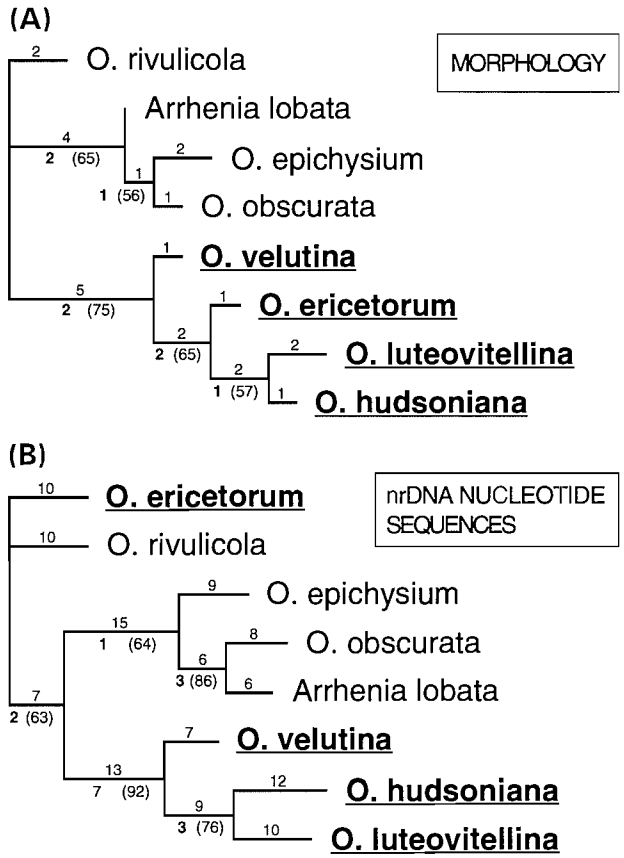
Results and discussion

Phylogenetic relationships among *Omphalina* species and *Arrhinia lobata*

The unweighted parsimony analysis of the morphological data matrix (Table 2) yielded one most parsimonious tree (Fig. 2A) of 38 steps (CI = 0.895, RI = 0.778, and RC = 0.696). The frequency distribution of tree lengths was skewed to the left with g1 = −0.63. As expected with such a small data set, which is often the case with morphological data sets, the internodes are not strongly supported by the data, with bootstrap values below 80% and decay values of two or one. It is worth noting, however, that the internode supporting the lichen clade is the most robust of the tree.

A single most parsimonious tree of 122 steps (Fig. 2B; CI = 0.639, RI = 0.506, RC = 0.323) was recovered with the unweighted parsimony analysis of the molecular data (Fig. 1). The histogram showing the frequency distribution of tree lengths was much more skewed to the left (g1 = −1.02) than for the morphological data. The nucleotide sequences strongly supported the *O. velutina* − *O. luteovitellina* − *O. hudsoniana* clade, with bootstrap values of 92% and a decay value of 7 (Fig. 2B). The *Arrhenia lobata* −

Fig. 2. (A) Single most parsimonious tree generated by an exhaustive search on 13 unweighted and unordered informative morphological characters (Table 2) using PAUP version 3.1.1 (Swofford 1993). (B) Single most parsimonious tree obtained from an exhaustive search on 62 unweighted informative and nonambiguously aligned nucleotide position using PAUP 3.1.1. For both trees, the lichenized species are in bold and underlined. Numbers of transformational changes (ACCTRAN optimization) are shown above each branch. Below each internode the decay value is shown in bold and the bootstrap value (1000 replications) is found in parentheses.

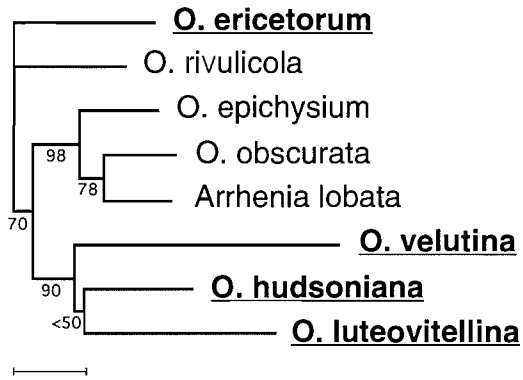


O. obscurata clade was also strongly supported by the molecular data. The nrDNA nucleotide sequences provided stronger support than the morphological data for *O. hudsoniana* being a sister species to *O. luteovitellina*.

To compare the effect of different evolutionary models on the reconstruction of the nucleotide sequence phylogeny, the neighbor-joining algorithm was implemented using the Jukes and Cantor (1969), the Kimura (1980), and the Kishino and Hasegawa (1989) models of nucleotide substitution. The same topology was recovered with these three different models (Fig. 3), which had no significant effect on the branch lengths. The same topology was recovered from the molecular data set with maximum likelihood, maximum parsimony, and neighbor-joining.

Only the morphological data suggested that the lichenized *Omphalina* species are monophyletic. The morphological and molecular trees in Fig. 2 differ only in the phylogenetic relationships of *O. epichysium* and *O. ericetorum*. An empirical comparison of three homogeneity tests is carried out below and the following three questions are addressed.

Fig. 3. Neighbor-joining tree generated using Jukes and Cantor's (1969) model of nucleotide substitution. The bootstrap values (1000 replications) are shown for each internode. The lichen-forming species are in bold and underlined. The scale is equal to 1 nucleotide substitution/100 sites. The topology of the most likely tree (ln likelihood = −2943.876 67, not shown) obtained using the Kishino and Hasegawa (1989) model of nucleotide substitution with DNAML was identical to the tree recovered using neighbor-joining and maximum parsimony.



(*i*) Which topology is the best estimate of the phylogenetic relationships of *Omphalina* species and *Arrhenia lobata*? (*ii*) Are the differences between the morphological tree and the molecular tree due to a different phylogenetic history between the species and the nrDNA sequences? (*iii*) Does combining the data sets provide a better estimate of the species phylogeny than either separate analysis?

Rodrigo et al. (1993) protocol for comparing morphological and molecular data sets

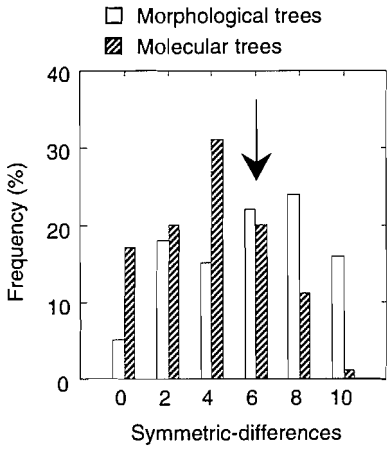
Is the tree based on morphology more similar to the tree based on nucleotide sequences than expected due to chance?

The first step in this procedure is the calculation of the pairwise symmetric-difference (Penny et al. 1982) between the morphological tree and the molecular tree. The pairwise distance, as implemented by PAUP's tree-to-tree distances menu command, was equal to 6. The smaller the value the more similar the trees are. The probability of obtaining a symmetric difference of 6 for eight taxa is 0.0048 (Hendy et al. 1984; Page 1989). Therefore, we can reject the null hypothesis at the 0.01 level that shared components of the two trees are due to chance. This test is especially important when estimating the phylogeny for a small number of taxa, because the probability of congruence among trees obtained from different data sets increases as the number of taxa decreases (Rodrigo et al. 1993; Page 1989).

Are the morphology and the molecular data sets the outcome of a common phylogenetic history?

To address this question Rodrigo et al. (1993) point out the need to first determine the variability associated with the morphological and molecular trees, respectively. This was performed by bootstrapping the morphological data set 1000 times followed by a branch-and-bound search on each of the 1000 bootstrapped data sets. All trees found were compiled by PAUP into one treefile. The same resampling and search

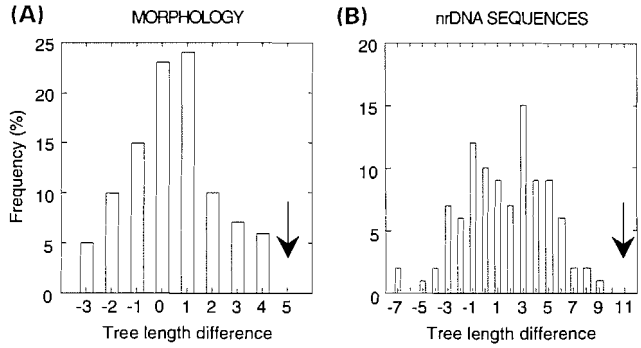
Fig. 4. Null distribution of pairwise symmetric-differences between most parsimonious trees obtained from two sets of 100 bootstrapped data sets where the trees obtained from the bootstrap analysis in both sets are estimating the same phylogeny. The open and hatched bars represent the null distribution for the morphological and molecular data sets, respectively. The arrow indicates the observed symmetric-difference between the most parsimonious morphological and molecular trees.



procedures were done on the molecular data set. A total of 7923 trees were obtained from the 1000 bootstrapped morphological data sets, whereas the molecular treefile contained 1774 trees resulting from this procedure. To determine if any of the trees in the morphological treefile shared the same topology as trees in the molecular treefile, the two treefiles were compared using reciprocally the mode 2 option of the GETTREES command in PAUP. A total of 88 trees were common to both the morphological and molecular tree files, representing nine different topologies. None of these nine topologies was identical to the most parsimonious trees obtained by analyzing both data sets separately (Fig. 2). When there is an overlap between the spectra of possible morphological and molecular trees, Rodrigo et al. (1993) suggest a supplementary method to test whether or not the most parsimonious morphological trees and molecular trees are sampling statistics of the same underlying phylogenetic history.

One way to test this is to ask whether the observed symmetric-difference of 6, measured between the morphological and molecular trees, is due to sampling error. To do this, Rodrigo et al. (1993) suggest generating a null distribution of tree distances given that the trees are estimating the same phylogeny. Their test consists of generating two sets of 100 bootstrap trees from one of the two original data sets, and then measuring the symmetric-difference for the 100 pairs of trees, for a total of 100 distances. This procedure is done separately on the morphological and molecular data sets. The frequency distribution of these distances represents the null distribution of distances when two trees estimate the same phylogenetic history. If more than 95% of the expected distances in the null distribution are smaller than the observed distance (=6), then the null hypothesis that the morphological and molecular trees are due to sampling error is rejected. When this test was applied to the morphological data sets (Fig. 4) only 38% of the expected distances in the null distribution were smaller than the observed distance (=6). Based

Fig. 5. (A) Null distribution using 100 permuted morphological data sets of pairwise tree length differences between the molecular tree and most parsimonious morphological tree (Fig. 2) used as constraints with PAUP. (B) Null distribution using 100 permuted molecular data sets of pairwise tree length differences between the most parsimonious morphological tree and molecular tree (Fig. 2) used as constraints with PAUP. The arrow indicates the observed length difference between the most parsimonious morphological and molecular trees.



on the null distribution of distances generated from the molecular data set, 68% of the expected distances were smaller than the observed distance (=6). So for both cases, we cannot reject the null hypothesis that the difference between the morphological tree and the molecular tree is due to sampling error. This supports a combined analysis of the two data sets.

T-PTP test

Another approach suggested by Bull et al. (1993) is to assume an a priori sampling model, as underlies the statistical tests that have been developed to compare support for two trees against a single data set. The Kishino and Hasegawa (1989) test as implemented in DNAML (PHYLIP 3.53c), and Faith's (1991) T-PTP test are examples of this approach. When the adapted version of the T-PTP test was applied reciprocally on the morphological and on the molecular data sets, none of the differences in length between the morphological and molecular trees was greater or equal to the observed difference of 5 or 11, respectively (Fig. 5). Since the observed differences in tree lengths between the morphological and molecular trees are significantly greater than expected, the T-PTP test would reject the null hypothesis ($P < 0.01$) that the two data sets in this study are homogeneous. Contrary to the Rodrigo et al. method (1993), this test would not support the fusion of these two data sets.

Likelihood test

This test developed by Kishino and Hasegawa (1989) as implemented in PHYLIP 3.53c can only be applied to the molecular data set. The question when applying this test becomes, is the most parsimonious morphological topology significantly worse in explaining the observed molecular data than the tree based on molecular data? The most parsimonious morphological tree was found to be significantly worse than the molecular tree to explain the observed sequences (Table 3). As with the T-PTP test, this test would suggest

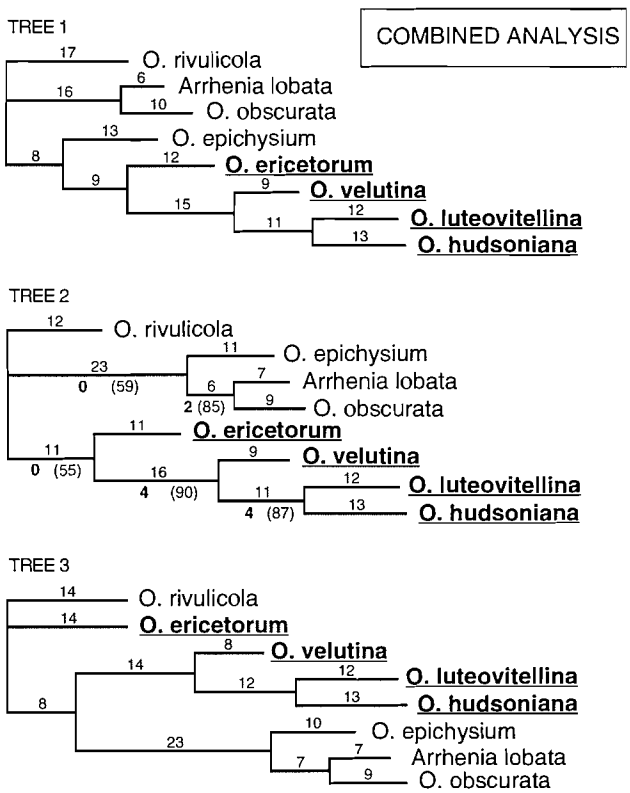
Table 3. Kishino and Hasegawa (1989) likelihood test as implemented in PHYLIP 3.53c (Felsenstein 1993).

Trees	Log likelihood	Difference of log likelihood	SD	Significantly worse?
Most parsimonious morphological tree	−2978.352 80	−34.476 13	12.9252	Yes
Molecular tree	−2943.876 85*			
Most parsimonious combined tree 1	−2959.284 60	−15.407 92	9.1620	No
Most parsimonious combined tree 2	−2951.951 36	−8.074 69	5.8111	No
Most parsimonious combined tree 3	−2943.876 85	−0.000 17	0.0252	No

NOTE: A given tree is considered significantly worse than the most likely tree in explaining the observed nucleotide sequences, when the difference of log likelihood is more than twice the standard deviation. The molecular tree represents the single topology that was recovered by maximum likelihood, maximum parsimony, and neighbor-joining.

*Best topology.

Fig. 6. Most parsimonious trees generated from an exhaustive search on a combined data set including 13 unweighted morphological characters and 62 unweighted nucleotide sites from the nrDNA large subunit. The lichenized species are in bold and underlined. Numbers of transformational changes are shown above each branch. Below each internode of tree number 2 the decay value is shown in bold and the bootstrap value (1000 replications) is found in parentheses.

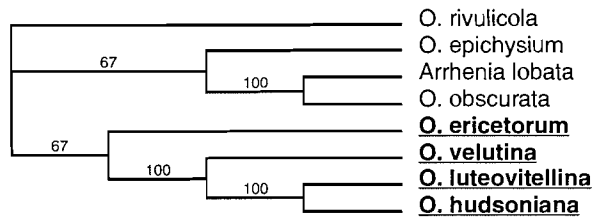


that the morphological and molecular data sets are not homogeneous and, therefore, should not be combined.

What is the best phylogenetic estimate for *Omphalina* species and *Arrhenia lobata*?

When two different data sets are samples of the same phylogenetic history, as was suggested by Rodrigo's test in this case study, it seems appropriate to combine the morphological and molecular data sets. Three most parsimonious trees of 165 steps (CI = 0.679, RI = 0.505, RC = 0.343) were

Fig. 7. Majority-rule consensus of the three most parsimonious trees (Fig. 6) revealed by the combined analysis.



generated by an exhaustive search on combined unweighted characters (Fig. 6). These three topologies only differ by the placement of the outgroup *O. rivulicola*. The single unrooted topology of the ingroup, as supported by the combined data set, is identical to the one revealed by the phylogenetic analysis of the nrDNA nucleotide sequences (Fig. 2B). This congruence suggests that this topology is our best estimate. The only conflicting evidence is the *O. epichysium* − *O. obscurata* clade of the most parsimonious morphological tree (Fig. 2A). This clade, however, is the most weakly supported by the morphological data, with a bootstrap value of 56%.

As to the phylogenetic relationship of the outgroup *O. rivulicola*, the morphological data supports rooting the ingroup on the same internode as the combined tree number 2 (Figs. 2A and 6). The molecular data suggest that the topology of the ingroup be rooted on the terminal branch leading to *O. ericetorum*, in agreement with the combined tree number 3 (Figs. 2B and 6). To summarize the different hypotheses for the rooting of the ingroup, a majority rule consensus of the combined three most parsimonious trees was done (Fig. 7). The combined trees 2 and 3 (Fig. 6) support the *O. epichysium* − *Arrhenia lobata* − *O. obscurata* monophyletic group, suggesting that the combined tree number one is less likely. The combined trees number 2 and 1 (Fig. 6) suggest that *O. ericetorum* forms a monophyletic group with the rest of the lichen-forming *Omphalina* species. This hypothesis is also supported by the morphological data (Fig. 2A), suggesting that the combined tree number 3 might be wrong. Consequently, the combined tree number 2 is our best estimate. This tree was found to be one of the 97 trees three steps longer than the most parsimonious morphological tree and one of the two trees two steps longer than the most parsimonious molecular tree.

Despite the fact that there were far more molecular characters (62) than morphological characters (13) in the combined data set, the molecular characters did not swamp

the morphological characters in determining the outcome of the phylogenetic analysis. For instance, the monophyly of the lichen-forming taxa obtained exclusively from the morphological evidence (Figs. 2A) was recovered in two of the three most parsimonious trees of the combined analysis (trees 1 and 2, Fig. 6). As expected, the bootstrap values from the combined data set were influenced almost entirely by the molecular data (Figs. 2 and 6). The decay values are less sensitive to the relative amount of characters coming from one or the other data sets.

Two main lines of evidence suggest that Rodrigo's test, as implemented on the data presented in this paper, is superior to the two other tests and was correct in suggesting that the two data sets are sampling the same phylogenetic history and could be combined. First, the g1 statistic increased slightly (−1.05) when the data sets were combined despite the fact that it was already very high for the molecular data alone (−1.02). Secondly, our best estimate (combined tree number 2, Fig. 6) is totally congruent with the relationships among these eight species in an analysis of the same portion of the nrDNA large subunit but extended to 26 species of *Omphalina* species and related genera (F. Lutzoni and R. Vilgalys, unpublished data). This suggests that as we gathered more molecular data, the phylogenetic analysis of the nucleotide sequences converged on one of the three most parsimonious trees (our best estimate, combined tree number 2; Fig. 6) obtained only when the smaller data sets were combined.

This difference in the outcome of the adapted T-PTP test and the Kishino and Hasegawa likelihood test versus Rodrigo's method can be explained by the nature of the questions addressed by each test. The Rodrigo et al. (1993) protocol addresses specific questions intrinsic to data set heterogeneity. The T-PTP test and the Kishino and Hasegawa likelihood test, as used here, are testing whether trees obtained from one data set are significantly different from trees based on another data set. This in itself is insufficient to determine whether two data sets are the outcome of a common phylogenetic history, since we still need to know if a significant difference between morphological and molecular trees is due to sampling error. The Rodrigo et al. (1993) protocol is the only method that explicitly addresses the specific question of sampling error, by using bootstrap replications. It seems more appropriate to use bootstrap (a resampling method) rather than permutation (where the taxon labels almost entirely lose their meaning) to estimate the variation associated with each data set. Rodrigo's protocol is also the only method that includes a test to determine whether molecular and morphological trees are more similar than pairs of random trees. This test is especially important when dealing with a small number of taxa, because as the number of taxa decreases, the probability that two trees from different data sets share the same components by chance alone increases.

Despite the greater potential of Rodrigo's method for testing data set homogeneity, we discovered one major weakness. One of the most critical steps in this method is to determine if the variation between topologies, obtained from the data sets when analyzed separately, can be explained by the variation associated with each data set. The criterion developed by Rodrigo et al. (1993) consists of showing that the spectra of possible topologies, obtained by bootstrapping the data sets separately, do not overlap. If this is demonstrated to be true, they suggest that we can reject the null hypothesis that both topologies estimate the same phylogeny. According to their criterion, only one shared tree between the two spectra of possible topologies needs to be found to accept the null hypothesis. The number of shared trees is dependent on the number of bootstrap replicates when dealing with small numbers of replicates. For instance, in this study, when we did only 100 bootstrap replicates, as in the example shown in the Rodrigo et al. (1993) paper, we did not find any trees in common. This was also the result in their study, which they interpreted as good evidence that the molecular and morphological signal did not share entirely the same historical information. Based on this decision, they proceeded to prune taxa from the two data sets, and reapplied the test. Our experience suggests that 100 bootstrap replicates is not sufficient. When we did 1000 bootstrap replicates, we found 88 shared trees representing nine topologies. This means that if we would do a series of 100 bootstrap replications, sometimes we would detect overlap and sometimes we would not.

Now the question is, does using a high number of bootstrap replicates bias the test toward accepting the null hypothesis that the two types of data are the result of the same underlying phylogenetic history? Our conclusions in this paper are based on the premise that the true organismal phylogeny of the eight species included in this study is the same as that shown by the analysis of the larger molecular data set (26 species). This premise may be incorrect and points to the need for simulation studies to explore the bias of Rodrigo's method associated with the number of bootstrap replicates.

Acknowledgments

We thank Dr. E.A. Zimmer for inviting F.L. to work at the Laboratory of Molecular Systematics, Smithsonian Institution, and providing access to an automated sequencer, as part of the Early Land Plant Evolution project. We are grateful to C.W. Cunningham, D. Hibbett, B. Mishler, and K.M. Pryer for their helpful criticisms of the manuscript. This research was supported in part by a Formation de chercheurs et l'aide à la recherche (FCAR) scholarship (Québec); an A.W. Mellon grant through the Plant Systematics Program at Duke University, a Smithsonian Mellon Junior Fellowship, a Mycological Society of America Graduate Fellowship, a New England Botanical Club Award for the Support of Botanical Research, a Sigma Xi Grant-in-Aid for research, the American Society of Plant Taxonomists, and an NSF Dissertation Improvement Grant (DEB 93-21828) to F.L.; and an NSF grant (BSR 88-06655) to R.V.

References

Alroy, J. 1994. Four permutation tests for the presence of phylogenetic structure. Syst. Biol. **43**: 430–437.

Archie, J.W. 1989. A randomization test for phylogenetic information in systematic data. Syst. Zool. **38**: 239–252.

Barrett, M., Donoghue, M.J., and Sober, E. 1991. Against consensus. Syst. Zool. **40**: 486–493.

Berbee, M.L., and Taylor, J.W. 1992*a*. Two Ascomycete

classes based on fruiting-body characters and ribosomal DNA sequences. Mol. Biol. Evol. **9**: 278–284.

Berbee, M.L., and Taylor, J.W. 1992*b*. 18S ribosomal RNA gene sequence characters place the human pathogen *Sporothrix schenckii* in the genus *Ophiostoma*. Exp. Mycol. **16**: 87–91.

Bruns, T.D., Vilgalys, R., Barns, S.M., Gonzalez, D., Hibbett, D.S., Lane, D.J., Simon, L., Stickel, S., Szaro, T.M., Weisburg, W.G., and Sogin, M.L. 1992. Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. Mol. Phylogenet. Evol. **1**: 231–241.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., and Waddell, P.J. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. **42**: 384–397.

Chippindale, P.T., and Wiens, J.J. 1994. Weighting, partitioning, and combining characters in phylogenetic analysis. Syst. Biol. **43**: 278–287.

de Queiroz, A. 1993. For consensus (sometimes). Syst. Biol. **42**: 368–372.

Doyle, J.J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Syst. Bot. **17**: 144–163.

Eernisse, D.J., and Kluge, A.G. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules and morphology. Mol. Biol. Evol. **10**: 1170–1195.

Faith, D.P. 1991. Cladistic permutation tests for monophyly and nonmonophyly. Syst. Zool. **40**: 366–375.

Faith, D.P., and Cranston, P.S. 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. Cladistics, **7**: 1–28.

Fell, J.W., Statzell-Tallman, A., Lutz, M., and Kurtzman, C.P. 1992. Partial rRNA sequences in marine yeasts: a model for identification of marine eukaryotes. Mol. Marine Biol. Biotechnol. **1**: 175–186.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. **27**: 401–410.

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

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.53c. University of Washington, Seattle, Wash.

Goldman, N. 1993. Statistical tests of models of DNA substitution. J. Mol. Evol. **36**: 182–198.

Gowan, S.P. 1989. A character analysis of the secondary products of the Porpidiaceae (lichenized Ascomycotina). Syst. Bot. **14**: 77–90.

Guadet, J., Julien, J., Lafay, J.F., and Brygoo, Y. 1989. Phylogeny of some *Fusarium* species, as determined by large subunit rRNA sequence comparison. Mol. Biol. Evol. **6**: 227–242.

Hendy, M.D., and Penny, D. 1982. Branch and bound algorithms to determine minimal evolutionary trees. Math. Biosci. **59**: 277–290.

Hendy, M.D., Little, H.C., and Penny, D. 1984. Comparing trees with pendant vertices labelled. SIAM J. Appl. Math. **44**: 1054–1065.

Hillis, D.M. 1987. Molecular versus morphological approaches to systematics. Annu. Rev. Ecol. Syst. **18**: 23–42.

Hillis, D.M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. *In* Phylogenetic analysis of DNA sequences. *Edited by* M.M. Miyamoto and J. Cracraft. Oxford University Press, New York. pp. 278–294.

Hillis, D.M. 1994. Homology in molecular biology. *In* Homology: the hierarchical basis of comparative biology. *Edited by* B.K. Hall. Academic Press, San Diego, Calif. pp. 339–368.

Hillis, D.M., and Huelsenbeck, J.P. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Hered. **83**: 189–195.

Huelsenbeck, J.P., Swofford, D.L., Cunningham, C.W., Bull, J.J., and Waddell, P.J. 1994. Is character weighting a panacea for the problem of data heterogeneity in phylogenetic analysis? Syst. Biol. **43**: 288–291.

Jukes, T.H., and Cantor, C.R. 1969. Evolution of protein molecules. *In* Mammalian protein metabolism. *Edited by* H. Munro. Academic Press, New York. pp. 21–132.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**: 111–120.

Kishino, H., and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. **29**: 170–179.

Kluge, A.G. 1989. A concern for evidence and a phylogenetic hypothesis of relationsips among *Epicrates* (Boidae, Serpentes). Syst. Zool. **38**: 7–25.

Korf, R.P. 1958. Japanese discomycete notes. Sci. Rep. Yokohama Natl. Univ. Sect. 2, Biol. Geol. Sci. **7**: 1–35.

Kurtzman, C.P., and Robnett, C.J. 1991. Phylogenetic relationships among species of *Saccharomyces*, *Schizosaccharomyces*, *Debaryomyces* and *Schwanniomyces* determined from partial ribosomal RNA sequences. Yeast, **7**: 61–72.

LoBuglio, K.F., and Taylor, J.W. 1993. Molecular phylogeny of *Talaromyces* and *Penicillium* species in subgenus *Biverticillium*. *In* The fungal holomorph: mitotic, meiotic and pleomorphic speciation in fungal systematics. Proceedings of an International Symposium, Newport, Oregon, August 4–7, 1992. *Edited by* D.R. Reynolds and J.W. Taylor. CAB International, Wallingford, England. pp. 115–119.

Lutzoni, F., and Vilgalys, R. 1995. *Omphalina* (Basidiomycota, Agaricales) as a model system for the study of coevolution in lichens. Cryptogam. Bot. **5**: 71–81.

Maddison, W.P., and Maddison, D.R. 1992. MacClade: analysis of phylogeny and character evolution. Version 3.0. Sinauer Associates, Sunderland, Mass.

Mishler, B.D., Donoghue, M.J., and Albert, V.A. 1991. The decay index as a measure of relative robustness within a cladogram. Hennig Society annual meeting. (Abstr.)

Miyamoto, M.M. 1985. Consensus cladograms and general classifications. Cladistics, **1**: 186–189.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nishida, H., and Sugiyama, J. 1993. Phylogenetic relationships among *Taphrina*, *Saitoella*, and other higher fungi. Mol. Biol. Evol. **10**: 431–436.

Page, R.D.M. 1989. COMPONENT user's manual (Version 1.5). University of Auckland, Auckland, New Zealand.

Penny, D., Foulds, L.R., and Hendy, M.D. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. Nature (London), **297**: 197–200.

Rehner, S.A., and Samuels, G.J. 1994. Taxonomy and phylogeny of *Gliocladium* analysed from nuclear large subunit ribosomal DNA sequences. Mycol. Res. **98**: 625–634.

Rodrigo, A.G., Kelly-Borges, M., Bergquist, P.R., and Bergquist, P.L. 1993. A randomisation test of the null hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. N.Z. J. Bot. **31**: 257–268.

Spatafora, J.W., and Blackwell, M. 1993*a*. The polyphyletic origins of ophiostomatoid fungi. Mycol. Res. **98**: 1–9.

Spatafora, J.W., and Blackwell, M. 1993*b*. Molecular systematics of unitunicate perithecial ascomycetes: the

Clavicepitales–Hypocreales connection. Mycologia, **85**: 912–922.

Swofford, D.L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? *In* Phylogenetic analysis of DNA sequences. *Edited by* M.M. Miyamoto and J. Cracraft. Oxford University Press, New York. pp. 295–333.

Swofford, D.L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Smithsonian Institution, Washington, D.C.

Tehler, A. 1988. A cladistic outline of the Eumycota. Cladistics, **4**: 227–277.

Wiley, E.O. 1981. Phylogenetics: the theory and practice of phylogenetic systematics. John Wiley & Sons, New York.