

# Phylogenetic analyses reveal deeply divergent species lineages in the genus *Sphaerobolus* (Phallales: Basidiomycota)

József Geml<sup>a,b,\*</sup>, Donald D. Davis<sup>a</sup>, David M. Geiser<sup>a,b</sup>

<sup>a</sup> Department of Plant Pathology, The Pennsylvania State University, University Park, PA, USA

<sup>b</sup> Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park, PA, USA

Received 14 January 2004; revised 22 September 2004

## Abstract

Phylogenetic analyses of 27 artillery fungus (*Sphaerobolus* sp.) isolates were conducted to identify species boundaries in the genus *Sphaerobolus*. Multiple gene genealogies inferred from maximum likelihood, Bayesian, and maximum-parsimony analyses of sequence data from individual loci (mtSSU, ITS, EF 1- $\alpha$ , and LSU) and a combined dataset (mtSSU, ITS, and EF 1- $\alpha$ ) concordantly indicate the existence of three deeply divergent lineages in the genus *Sphaerobolus*, each representing a phylogenetic species. These three phylogenetic species correspond to two known species: *Sphaerobolus iowensis* and *Sphaerobolus stellatus*, and a newly discovered species. Suprageneric phylogenetic analyses of the mtSSU and LSU datasets containing representatives of related genera of the gomphoid–phalloid clade of Homobasidiomycetes suggested that the undescribed taxon likely is more closely related to *S. stellatus* than to *S. iowensis*.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** *Sphaerobolus*; Phylogenetic species recognition; Evolution; Basidiomycota

## 1. Introduction

Since the first documentation of *Sphaerobolus* nearly 300 years ago under the name *Carpobolus* (Micheli, 1729), many mycologists have been fascinated by this unique fungus. Its common name “artillery fungus” comes from its capability of ejecting a 1-mm diameter “gleba” (spore mass) up to 6 m toward the brightest light in its environment (Buller, 1933; Walker, 1927). The first species in the genus was *S. stellatus* (Tode) Pers. Walker (1927) later described a second species (*S. iowensis*) and a new variety (*S. stellatus* var. *giganteus*). These fungi are extremely common in temperate climates, encountered most commonly on wood mulches used in landscaping. Several

researchers since have studied the growth and reproduction of the artillery fungus (Alasoadura, 1963; Buller, 1933; Dykstra, 1982; Fletcher and Cooke, 1984; Ingold, 1972; Walker, 1927; Walker and Anderson, 1925). In recent years, the artillery fungus has become a source of distress to homeowners, landscape mulch producers, and insurance companies due to the strong adhesion of the discharged gleba to artificial surfaces including house siding, cars, and windows (Lehman, 1985). Because of these problems, recent attention has been given to potential control measures inhibiting the growth of the fungus (Brantley et al., 2001a,b).

Based on morphology, the genus *Sphaerobolus* has been classified as a member of the class Gasteromycetes along with other fungi having passive spore discharge, including bird’s nest fungi, puffballs, earth balls, stink-horns, and earth stars. Within Gasteromycetes, authors have placed the artillery fungus into different families and orders, including Sphaerobolaceae in Nidulariales

\* Corresponding author. Institute of Arctic Biology, 311 Irving I Building, 902 N. Koyukuk Drive, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA. Fax: +1 907 474 6967.

E-mail address: [jgeml@iab.alaska.edu](mailto:jgeml@iab.alaska.edu) (J. Geml).

(Ulloa and Hanlin, 2000) and Sphaerobolaceae in Sclerodermatales (Hawksworth et al., 1996). However, molecular phylogenetic analyses of Hibbett et al. (1997) revealed Gasteromycetes to be a polyphyletic taxon that does not represent a true evolutionary group within the Basidiomycota. Sequences of nuclear and mitochondrial genes placed the genus *Sphaerobolus* in the gomphoid–phalloid clade with an extremely morphologically diverse set of genera as closest relatives, including *Geastrum*, *Phallus*, *Pseudococcus*, *Ramaria*, *Clavariadelphus*, *Gomphus*, and *Gautieria* (Hibbett et al., 1997; Moncalvo et al., 2002). These results are incorporated in the most recent classification by Kirk et al. (2001), who placed *Sphaerobolus* in the family Geastraceae, order Phallales, class Basidiomycetes, phylum Basidiomycota.

The two currently recognized species, *S. stellatus* (Tode) Pers. and *S. iowensis* Walker (Hawksworth et al., 1996), are distinguished by micromorphological characteristics (Walker, 1927). Early gleba development in young basidiocarps of *S. stellatus* is marked by the formation of globular knots of binucleate hyphae. These later become centers around which new basidia are formed, eventually producing four (occasionally eight) basidiospores per basidium. The basidia entirely break down and disappear as soon as the spores are mature, making room for the enlargement and maturation of other basidia. In *S. iowensis*, the development of young basidia is followed by the formation of characteristic cavities or chambers in which the spores are produced (Walker, 1927). In addition, the gleba covering (peridium) consists of a three-layered endoperidium and three-layered exoperidium in *S. stellatus*, whereas in *S. iowensis* a gelatinous layer in the outer peridium is absent (Walker, 1927). Several other names of unknown origin exist in herbaria, including *S. bombardioides*, *S. carpobolus*, *S. corii*, *S. crustaceus*, *S. epigaeus*, *S. minimus*, *S. minutissimus*, *S. rubidus*, *S. sparsus*, and *S. tubulosus*. The origin of these binomials cannot be found in the literature, and thus cannot be considered scientifically valid.

Since no molecular systematic work had been published investigating species limits in this genus, our goal was to elucidate the molecular phylogenetics of *Sphaerobolus* species to provide better understanding of the biology of the artillery fungus. For this purpose we generated nucleotide sequences of three ribosomal DNA regions and one protein coding gene region: mitochondrial ribosomal RNA small subunit (mtSSU), internal transcribed spacer regions of the nuclear ribosomal gene repeat (ITS), the nuclear large ribosomal RNA subunit (LSU), and translation elongation factor 1- $\alpha$  (EF 1- $\alpha$ ). These regions have been successfully used in several studies to reveal both deep and shallow phylogenies in fungi (Baayen et al., 2001; Bruns et al., 1991; Hillis and Dixon, 1991; Hopple and Vilgalys, 1999; Lutzoni and Vilgalys, 1995; Moncalvo et al., 1995; O'Donnell et al., 1998; 2001; Pine et al., 1999) and are well suited for

determining phylogenetic species boundaries using genealogical concordance as outlined by Taylor et al. (2000).

## 2. Materials and methods

### 2.1. Isolates and DNA extraction

Isolates of the artillery fungus included in this study either came from culture collections (public or private) or were collected as gleba from landscape mulch and surrounding objects (Table 1). To isolate pure cultures, the collected gleba were surface-sterilized by agitation in 20% bleach solution for 3 min, washed with distilled water, air-dried, sprayed with 70% ethanol, and finally air-dried again on filter paper under aseptic conditions. The surface-sterilized gleba were placed on DIFCO oatmeal agar (OA; Becton–Dickinson Microbiology Systems, Sparks, MD) and grown at room temperature. All isolates are available by request from the PSU Mushroom Culture Collection. Mycelia from each isolate were grown at room temperature in liquid Difco potato dextrose broth, harvested, ground, and DNA was isolated using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA).

### 2.2. PCR amplification and DNA sequencing

Portions of the mitochondrial small ribosomal subunit gene (mtSSU), nuclear large ribosomal subunit gene (LSU), translation elongation factor 1- $\alpha$  gene (EF 1- $\alpha$ ), and the entire nuclear internal transcribed spacer (ITS) + 5.8S ribosomal subunit gene region were amplified in PCR mixtures containing 37  $\mu$ l PCR water, 5  $\mu$ l 10 $\times$  PCR buffer (0.5 M KCl, 0.1 M Tris–HCl, pH 8.3, and 0.025 M MgCl<sub>2</sub>), 5  $\mu$ l 10 $\times$  dNTPs (2 mM of each dNTP), 0.1  $\mu$ l AmpliTaq DNA polymerase (Perkin-Elmer, Foster City, CA), and 1  $\mu$ l of 10  $\mu$ M forward primer and reverse primer for the region of interest, and 1  $\mu$ l template DNA (100-fold dilution of original DNA solution extracted). PCRs were performed in a 96-well thermocycler (PTC-100 Programmable Thermal Controller, MJ Research) using the following temperature program for the three ribosomal gene regions: 94 °C/5 min; 34 cycles of 94 °C/1 min, 53 °C/1 min, 72 °C/1 min; and 72 °C/5 min. However, for the single copy gene EF 1- $\alpha$  a “touchdown” PCR setting was used with an annealing temperature of 65 °C in the first cycle, then successively reduced by 1 °C per cycle to 56 °C, after which the annealing temperature was maintained at 56 °C for the remaining 30–36 cycles (Stephen Rehner, pers. comm.). The following primers were used for amplification: ITS5 and ITS4 for ITS (White et al., 1990), NL1 and NL4 for LSU (O'Donnell, 1996), MS1 and MS2 for mtSSU (White et al., 1990), and EF1-983F (GCY CCY GGH CAY CGT GAY TTY AT) and EF1-1567R (ACH

Table 1  
*Sphaerobolus* isolates investigated in this study

Isolate code	Origin	GenBank Accession No.			
		mtSSU	ITS	EF 1- $\alpha$	LSU
SS-1	IN, collected from mulch	AY488000	AY487950	AY487976	—
SS-2	Elizabethtown, PA, collected from mulch	AY488001	AY487951	AY487977	—
SS-3	State College, PA, collected from mulch	AY488002	AY487952	AY487978	—
SS-4	Langhorne, PA, collected from mulch	AY488003	AY487953	AY487979	—
SS-5	PA, collected from mulch	AY488004	AY487954	AY487980	—
SS-7	West Mifflin, PA, collected from mulch	AY488005	AY487955	AY487981	—
SS-9	Chapel Hill, NC, collected from mulch	AY488006	AY487956	AY487982	AY439010
SS-10	ATCC 18339 ( <i>S. stellatus</i> ), American Type Culture Collection, MD	AY488007	AY487957	AY487983	AY439011
SS-11	ATCC 52850 ( <i>S. iowensis</i> ), American Type Culture Collection, East Lansing, MI	AY488008	AY487958	AY487984	AY439014
SS-12	DSH 96-015 ( <i>S. stellatus</i> ), Great Brook State Park, MA	AY488009	AY487959	AY487985	—
SS-13	Erie, PA, collected from mulch	AY488010	AY487960	AY487986	—
SS-14	Lucinda, PA, collected from mulch	AY488011	AY487961	AY487987	—
SS-16	Olney, MD, collected from mulch	AY488012	AY487962	AY487988	—
SS-17	Olney, MD, collected from mulch	AY488013	AY487963	—	—
SS-18	Olney, MD, collected from mulch	AY488014	AY487964	AY487989	—
SS-19	Atlanta, GA; culture received	AY488015	AY487965	AY487990	AY439012
SS-20	Olney, MD, collected from mulch	AY488016	AY487966	AY487991	—
SS-21	Galion, OH, collected from mulch	AY488017	AY487967	AY487992	—
SS-22	Ithaca, NY, collected from mulch	AY488018	AY487968	AY487993	—
SS-23	Medina, OH, collected from mulch	AY488019	AY487969	AY487994	—
SS-24	Russell, PA, collected from mulch	AY488020	—	—	—
SS-25	Newton Centre, MA, collected from mulch	AY488021	AY487970	AY487995	—
SS-26	# 9597, Inst. for Fermentation, Otsu, Japan	AY488022	AY487971	AY487996	AY439013
SS-27	Anchorage, AK, collected from mulch	AY488023	AY487972	—	—
SS-28	Anchorage, AK, collected from mulch	AY488024	AY487973	AY487997	—
SS-29	# MIN 864513 ( <i>S. stellatus</i> ), Bell Museum of Nat. Hist., Elm Creek Nature Reserve, MN	AY488025	AY487974	AY487998	—
SS-31	CBS#321.32 ( <i>S. stellatus</i> ), Centraalbureau voor Schimmelcultures, The Netherlands	AY488026	AY487975	AY487999	—

Since identification of isolates collected from gleba could not be made with full confidence, species names are mentioned only for culture collection isolates.

GTR CCR ATA CCA CCR ATCTT) for EF 1- $\alpha$  (Stephen Rehner, pers. comm.). Amplification products were electrophoresed in a 3.0% agarose gel and stained with ethidium bromide for visualization of the bands. PCR products were purified directly using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA).

Purified amplification products were sequenced using the Applied Biosystems (ABI) BigDye v. 3.0 terminator kit and an ABI 377 automated DNA sequencer (Perkin-Elmer, Foster City, CA). Each sample was sequenced in both directions with the same primers used for PCR. The only exception was the 1567R primer of the EF 1- $\alpha$  gene that was replaced by the 1567Ra (ACH GTR CCR ATA CCA CC) primer in sequencing for better results (Stephen Rehner, pers. comm.).

### 2.3. Intra-genetic phylogenetic analyses

Sequence data obtained for both strands of each locus were edited and assembled for each isolate using Sequencher 3.1 (Gene Codes, Ann Arbor, MI). Sequence alignments were initiated using Clustal X (Thompson

et al., 1997) and subsequently corrected manually. Analyses were conducted in multiple steps using maximum-likelihood (ML) and maximum-parsimony (MP) methods in PAUP\* 4b10 (Swofford, 2002), and Bayesian analysis in Mr.Bayes 3.0 (Huelsenbeck and Ronquist, 2001). Since these methods follow different theories and algorithms, congruent features found in all three types of analyses were considered meaningful.

mtSSU, ITS, and EF 1- $\alpha$  sequences generated from 27, 26, and 24 *Sphaerobolus* isolates, respectively, were used for intra-genetic phylogenetic analyses. Ambiguously aligned regions were excluded in the ML and Bayesian analyses, consisting of the following positions: 47–63, 205–255, and 296–308 in the intra-genetic mtSSU, 79–84, 160–221, 487–499, 623–666, and 699–703 in the ITS, and 203–260, 408–460, and 558–562 in the EF 1- $\alpha$  alignments. Interestingly, regions that could not be aligned unambiguously between the three different phylogenetic species were highly similar and sometimes identical within lineages, indicating the existence of fixed differences. Different evolutionary models with varying values of base frequencies, substitution types,

$\alpha$ -parameter of the  $\gamma$ -distribution of variable sites, and proportion of invariable sites, among other parameters, were compared via the likelihood ratio test for each locus using PAUP\* and Modeltest 3.06 (Posada and Crandall, 1998) to determine the best-fit evolutionary model for both ML and Bayesian analyses. ML analyses were carried out with the heuristic search option using the “tree bisection and reconnection” (TBR) algorithm with 100 random sequence additions to find the overall optimum instead of local optima. To test the statistical reliability of the generated trees and test the stability of clades, the bootstrap test (Felsenstein, 1985) was used with “full heuristic search” and 100 replicates. In Bayesian phylogenetic analyses, 100,000 generations were run in four chains. The chains were sampled every 100th generation. When the likelihood scores of trees sampled approached similar values, they were considered to have converged. In each run, only trees after this convergence point were included in computing the consensus tree.

In MP analyses, previously excluded ambiguous regions were included after being recoded with programs INAASE 2.3b (Lutzoni et al., 2000) and ARC (Ambiguous Regions Coding) v.1.5 (Miadlikowska et al., 2003). Ambiguous regions shorter than 20 bp were recoded using INAASE, while regions equal to or longer than 20 nucleotides were recoded with ARC. The code matrices obtained via these programs were attached to the appropriate alignments and were included in MP analyses. A heuristic search was carried out with the “tree bisection and reconnection” (TBR) algorithm with 100 random addition sequences to find the overall optimum instead of local optima. The bootstrap test was used with 500 replicates.

Phylogenetic trees obtained by analyzing each locus were compared to detect clades that are supported by every tree. Phylogenetic species were recognized based on the criteria of Taylor et al. (2000) by determining the transition points between concordant and conflicting information derived from different gene genealogies. Gene genealogies should be congruent between phylogenetic species, where gene trees represent organismal trees. In contrast, gene genealogies should be incongruent within species as the result of recombination of unlinked loci by independent assortment of chromosomes or by crossing over during sexual reproduction. Only clades supported by all three loci and all analyses were considered as phylogenetic species.

#### 2.4. Suprageneric phylogenetic analyses

To test the monophyly of the genus *Sphaerobolus* in suprageneric phylogenetic analyses, two alignments (consisting of mtSSU and LSU sequences) were made including at least two representatives of the three putative phylogenetic species in the genus *Sphaerobolus*

as inferred in our earlier analyses. The alignments contained all available artillery fungus sequences in GenBank: AF026662 (mtSSU), AF139975, and AF393077 (LSU), and homologous sequences of other genera from the gomphoid–phalloid clade (mtSSU and LSU, respectively): *Clavariadelphus* (AF185972, AF213133), *Gautieria* (AF213144, AF336249), *Geastrum* (AF026685, AF287859), *Pseudocolus* (AF026666, AF518641), *Ramaria* (AF213138, AF213114). The abundance of unalignable regions in ITS and EF 1- $\alpha$ , and the scarcity of EF 1- $\alpha$  sequences from *Sphaerobolus* and related genera in GenBank made these two genes unsuitable for suprageneric analyses. Ambiguously aligned regions were excluded in the ML and Bayesian analyses, consisting of the following positions: 208–259, 299–309, 324–330, 338–342, and 470–480 in the suprageneric mtSSU. Determination of the best-fit evolutionary models and the ML, Bayesian, and MP analyses were carried out as described earlier.

#### 2.5. Polymorphism and divergence

The number of polymorphic sites and their distribution among species was determined for sequence data generated from three loci (mtSSU, ITS, and EF 1- $\alpha$ ). LSU was not included in this part of the analyses due to the low level of variability in this gene region and the small number of isolates sequenced. Within species, nucleotide variability was measured using  $\theta$ , the proportion of polymorphic sites in a sample (Watterson, 1975), and  $\pi$ , the average number of nucleotide differences among sequences in a sample (Nei and Li, 1979). Tajima's  $D$  (Tajima, 1989) and Fu and Li's  $D^*$  and  $F^*$  (Fu and Li, 1993) test statistics were calculated to test for departures from the neutral theory of molecular evolution (Kimura, 1983). The neutral theory predicts that  $\theta$  and  $\pi$  (compared by Tajima's  $D$ ) or  $h$  and  $S$  (the number of singleton mutations and the number of segregating sites, compared by Fu and Li's statistics) should estimate the same parameter,  $4N_e\mu$ , under the neutral model; therefore, significant differences in these estimates may indicate departure from neutrality. Measures of variation, calculation of the estimated minimum number of recombination events ( $R_m$ ), and the tests for neutrality were performed with the computer program DnaSP v. 3.51 (Rozas and Rozas, 1999).

### 3. Results

#### 3.1. Intra-generic phylogenetic analyses

The intra-generic mtSSU, ITS, EF 1- $\alpha$ , and the combined datasets consisted of 583, 757, 624, and 1964 characters, respectively, including gaps. Of these 81, 130, 116, and 323 ambiguous positions were excluded in the intra-generic analyses. Three types of long (>1 kb), indepen-

dent insertions, each with a unique insertion site, length, and nucleotide sequence, were detected in the mtSSU region of some of the isolates. These insertions were excluded from the alignment, but were considered as additional characters in MP analyses. The three insertion types were included as three characters, each coded as 0 (absent) versus 1 (present). The Hasegawa–Kishino–Yano model (Hasegawa et al., 1985) with no proportion of invariable sites ( $I=0$ ) and equal variation rates for all sites (HKY) was selected by hierarchical likelihood ratio tests as the best-fit evolutionary model for the mtSSU and ITS datasets. The base frequencies for the mtSSU and ITS alignments were  $\text{freqA}=0.3304$ ,  $\text{freqC}=0.1610$ ,  $\text{freqG}=0.2410$ ,  $\text{freqT}=0.2676$ , and  $\text{freqA}=0.2617$ ,  $\text{freqC}=0.2256$ ,  $\text{freqG}=0.2012$ ,  $\text{freqT}=0.3115$ , respectively. The transition/transversion ratios were  $\text{ti/tv}=1.4824$  and  $\text{ti/tv}=1.8035$ . The Tamura–Nei model (Tamura and Nei, 1993) with no proportion of invariable sites ( $I=0$ ) and estimated  $\alpha$ -parameter of  $\gamma$ -distribution (TN + G) was selected as the best-fit evolutionary model for the EF 1- $\alpha$  dataset. The base frequencies were  $\text{freqA}=0.2349$ ,  $\text{freqC}=0.2869$ ,  $\text{freqG}=0.2492$ ,  $\text{freqT}=0.2290$ . The substitution rate matrix was as follows:  $R(a)$  [A–C]=1.0000,  $R(b)$  [A–G]=2.7043,  $R(c)$  [A–T]=1.0000,  $R(d)$  [C–G]=1.0000,  $R(e)$  [C–T]=11.3650,  $R(f)$  [G–T]=1.0000. The among-site rate variation was characterized by variable sites following a  $\gamma$ -distribution with an estimated  $\alpha$ -parameter=0.3767. Since the HKY model was determined earlier as best-fit model for two of the three loci, that model was chosen for ML and Bayesian analyses of the combined dataset as well.

Likelihood values converged after about 16200, 24000, 22000, and 16100 generations in Bayesian analysis of the mtSSU, ITS, EF 1- $\alpha$ , and combined datasets, respectively. The consensus trees were computed from 839, 761, 781, and 840 trees, after discarding the first 162, 240, 220, and 161 trees as “burn in.” The Bayesian trees of individual loci and the combined dataset are shown in Figs. 1A, C, and E and Fig. 2A, respectively, with bootstrap (ML) and posterior probability (Bayesian) values of the supported branches. After including the character matrices of the ambiguous regions recoded by INAASE and ARC, the final alignments of mtSSU, ITS, and EF 1- $\alpha$  for MP analyses consisted of 530, 676, 555, and 1761 characters, respectively. The single most parsimonious tree (mtSSU), and one of the 6 (ITS), 122 (EF 1- $\alpha$ ), and 90 (combined) most parsimonious trees are shown with bootstrap values in Figs. 1B, D, and F and Fig. 2B, respectively.

In all analyses, *Sphaerobolus* isolates formed three deeply divergent, highly supported clades. One of the clades corresponds to *S. stellatus*, as indicated by the presence of two isolates previously identified as *S. stellatus* from different culture collections: ATCC 18339, CBS 321.32. Isolates in the second clade are considered *S. iowensis*, as this group included a specimen previously identified as *S. iowensis* (ATCC 52850). The third clade

has no taxonomic connection and represents an undescribed or unknown taxon. The species-level clades received 100% support in all analyses, and numerous subgroups receiving varying levels of support were found in *S. iowensis* and in *S. stellatus*.

### 3.2. Suprageneric phylogenetic analyses

The mtSSU and LSU alignments of *Sphaerobolus* and related genera consisted of 601 and 577 characters, respectively, including gaps. Eighty-six ambiguous positions were excluded from the mtSSU dataset, while the entire LSU region aligned well. The Hasegawa–Kishino–Yano model with no proportion of invariable sites ( $I=0$ ) and estimated  $\alpha$ -parameter of  $\gamma$ -distribution (HKY + G) and the Tamura–Nei model with equal base frequencies, no proportion of invariable sites ( $I=0$ ) and estimated  $\alpha$ -parameter of  $\gamma$ -distribution (TNef + G) were selected as the best-fit evolutionary models for the mtSSU and LSU datasets, respectively. Likelihood values of Bayesian analysis converged after about 9100 and 10200 generations in mtSSU and LSU datasets, respectively, and the consensus tree was computed from 910 and 899 trees after discarding the first 91 and 102 trees as “burn in.” After including the character matrix of the ambiguous regions recoded by INAASE and ARC, the final alignment of mtSSU for MP analyses consisted of 543 characters. MP bootstrap cladograms for both datasets are shown in Fig. 3. ML, MP, and Bayesian trees showed all *Sphaerobolus* isolates forming a monophyletic group, with strong bootstrap and posterior probability values in the mtSSU (88% [ML], 91% [MP], and 100% [Bayesian]). However, in the LSU data, the monophyly of all *Sphaerobolus* isolates received weak support in all three analyses, the MP analysis giving the highest value of 60%. Within the genus *Sphaerobolus*, all of the three phylogenetic species received 100% support (ML, MP, and Bayesian) in the mtSSU analyses, while clades in the LSU tree, representing *S. iowensis*, *S. stellatus* and the undescribed taxon, were supported by bootstrap and posterior probability values of 88, 99, and 96% (ML), 97, 100, and 99% (MP) and all 100% (Bayesian), respectively. Furthermore, *S. stellatus* and the undescribed *Sphaerobolus* taxon formed a subgroup with strong and moderate supports in the mtSSU (83% [ML], 94% [MP], and 99% [Bayesian]) and LSU (<50% [ML], 60% [MP], and 81% [Bayesian]) trees. One of the *Sphaerobolus* LSU sequences from GenBank (AF139975) was nested in the clade representing the undescribed taxon, providing independent evidence for the existence of this group.

### 3.3. Polymorphism and divergence

*Sphaerobolus* sequences of mtSSU, ITS, and EF 1- $\alpha$  varied between 538–569, 677–713, and 557–571 bp in length, respectively, not including insertions in the

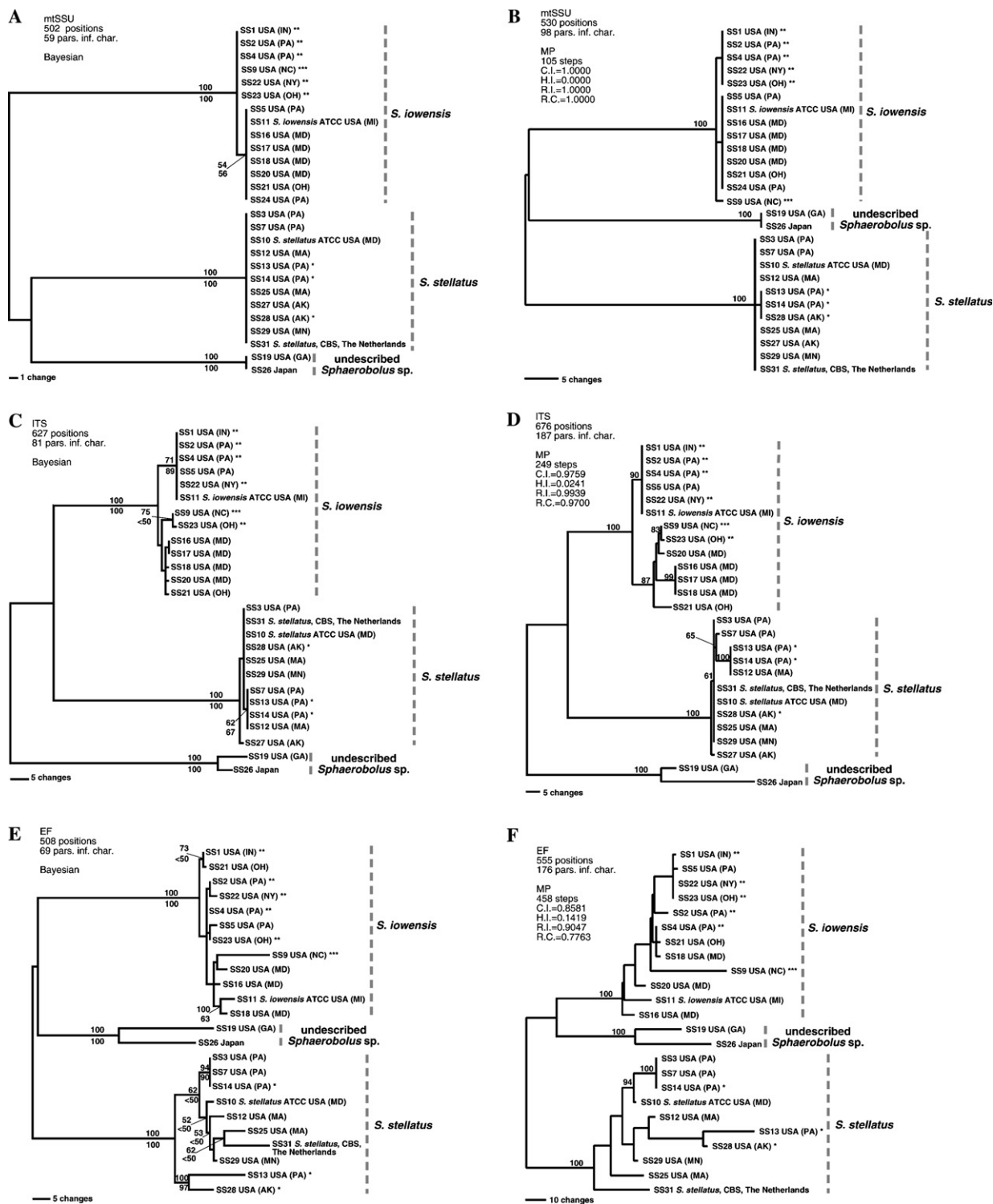


Fig. 1. Phylograms of *Sphaerobolus* species based on Bayesian (A, C, and E) and maximum parsimony analyses (B, D, and F) of the mtSSU (A and B), ITS (C and D) and EF 1- $\alpha$  (E and F) datasets. The trees are midpoint rooted. Posterior probabilities from Bayesian and bootstrap values from maximum parsimony analyses are indicated in percents above and below the branches on Bayesian phylograms. Bootstrap values of maximum parsimony analyses are shown on the most parsimonious trees. Symbols \* and \*\* represents mtSSU insertion Types A, B, and C, respectively.

mtSSU (Table 2). There were eight and one ITS and nine and 24 EF 1- $\alpha$  parsimony informative sites, respectively, in the *S. iowensis* and *S. stellatus* intraspecific datasets. In the intraspecific mtSSU datasets, only a single parsim-

ony-informative nucleotide substitution was identified in *S. iowensis*, and none in the other two lineages. However, as discussed earlier, insertion types within the mtSSU dataset were parsimony informative. The total

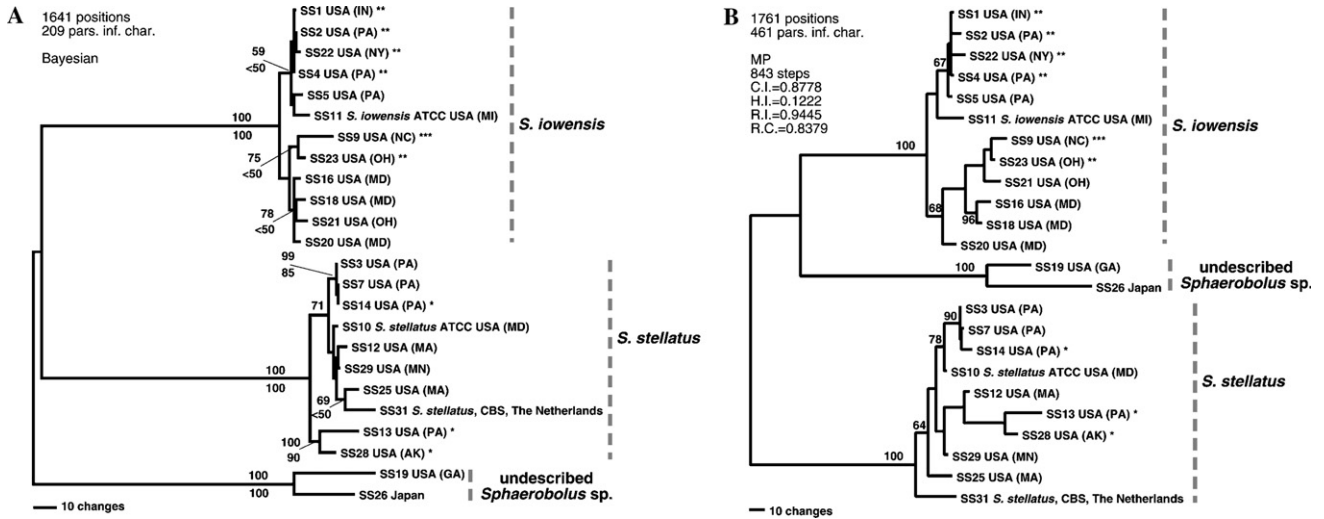


Fig. 2. Phylogenies of *Sphaerobolus* species based on Bayesian (A) and maximum parsimony (B) analyses of the combined (mtSSU, ITS, and EF 1- $\alpha$ ) dataset. The trees are midpoint rooted. Posterior probabilities from Bayesian and bootstrap values from maximum likelihood analyses are indicated in percents above and below the branches on Bayesian phylogenies. Bootstrap values of maximum parsimony analyses are shown on the most parsimonious trees. Symbols \*, \*\*, and \*\*\* represents mtSSU insertion Types A, B, and C, respectively.

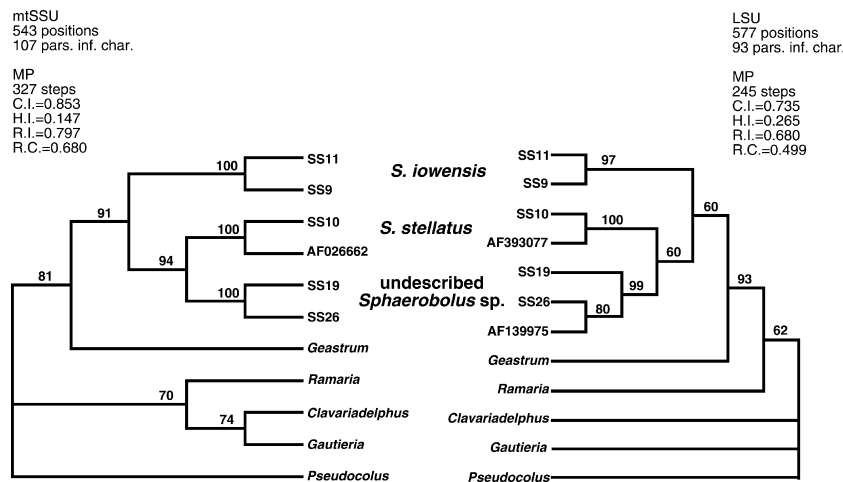


Fig. 3. Bootstrap cladograms of *Sphaerobolus* and related genera based on maximum parsimony analysis of the mtSSU and LSU rDNA datasets. The trees are rooted with *Pseudocolus* as outgroup. Bootstrap values are indicated above the branches.

number of nucleotide substitutions, including autapomorphies, is shown in Table 2.

At most of the loci within species, values of  $\pi$  and  $\theta$  were similar, and neither Tajima's  $D$  nor Fu and Li's  $D^*$  and  $F^*$  differed significantly from zero, as expected under the neutral theory. The only exception was the EF 1- $\alpha$  locus of *S. iowensis* which showed a significantly negative Tajima's  $D$ , and near-significant ( $0.1 > P > 0.05$ ) Fu and Li's  $D^*$  and  $F^*$  values. Although the only non-zero  $R_m$  value ( $R_m = 7$ ) was observed with the highest  $S$  value ( $S = 47$ ) in *S. stellatus*, all other loci showed no evidence for intralocus recombination in spite of having relatively high proportions of segregating sites (Table 2).

Interspecific divergence is summarized in Table 3. Pairwise divergence values between all three taxa were similarly high, with a large number of fixed differences, low numbers of polymorphisms shared among taxa, and

a varying number of sites that were polymorphic in one species but monomorphic in the other. As expected from the intraspecific polymorphism values discussed earlier, EF 1- $\alpha$  contained the highest number of sites polymorphic in one species but monomorphic in the other, followed by ITS, and mtSSU. Polymorphism to divergence ratios varied greatly between loci and species comparisons from 0.0063 in mtSSU (*iow-ste*) to 0.24101 in EF 1- $\alpha$  (*ste-und*).

#### 4. Discussion

Data from gene genealogies from each locus, and the polymorphism and divergence analyses concordantly suggested three deeply divergent lineages corresponding to the taxa *S. iowensis*, *S. stellatus* and a newly discov-

Table 2  
Nucleotide polymorphism in *S. iowensis*, *S. stellatus*, and the undescribed *Sphaerobolus* taxon

Locus/species	No. of sequences	Total sites	Segregating sites	$\pi$ (per site)	$\theta$ (per site)	Tajima's <i>D</i>	Fu and Li's <i>D</i> *	Fu and Li's <i>F</i> *	<i>R<sub>m</sub></i>
mtSSU									
<i>S. iowensis</i>	12	538	1	0.00099 ( $<1 \times 10^{-7}$ )	0.00062 ( $4 \times 10^{-7}$ )	1.38110	0.75202	1.02055	0
<i>S. stellatus</i>	11	569	0	0 (0)	0 (0)	—	—	—	0
Undescribed	2	546	0	0 (0)	0 (0)	N/A <sup>a</sup>	N/A <sup>a</sup>	N/A <sup>a</sup>	0
ITS									
<i>S. iowensis</i>	12	677	9	0.00430 ( $3 \times 10^{-7}$ )	0.00440 ( $4.7 \times 10^{-6}$ )	−0.09692	−0.32484	−0.30209	0
<i>S. stellatus</i>	11	713	3	0.00122 ( $2 \times 10^{-7}$ )	0.00144 ( $9 \times 10^{-7}$ )	−0.50634	−0.87363	−0.88004	0
Undescribed	2	682	10	0.01466 ( $5.4 \times 10^{-5}$ )	0.01466 ( $1.2 \times 10^{-4}$ )	N/A <sup>a</sup>	N/A <sup>a</sup>	N/A <sup>a</sup>	0
EF 1- $\alpha$									
<i>S. iowensis</i>	12	557	19	0.00642 ( $5.4 \times 10^{-6}$ )	0.01130 ( $2.4 \times 10^{-5}$ )	−1.88915 <sup>c</sup>	−2.05526 <sup>b</sup>	−2.28979 <sup>b</sup>	0
<i>S. stellatus</i>	10	557	47	0.02553 ( $3 \times 10^{-5}$ )	0.02983 ( $1.6 \times 10^{-4}$ )	−0.78982	−0.70305	−0.81849	7
Undescribed	2	571	42	0.07356 ( $1.4 \times 10^{-3}$ )	0.07356 ( $2.8 \times 10^{-3}$ )	N/A <sup>a</sup>	N/A <sup>a</sup>	N/A <sup>a</sup>	0

Values of the average number of nucleotide differences ( $\pi$ ) and the proportion of polymorphic sites ( $\theta$ ) are given per site with variance estimates in parentheses. Tajima's *D*, Fu and Li's *D*\* and *F*\* statistics were calculated to test departure from neutrality. *R<sub>m</sub>* is the minimum number of recombination events per locus based on the presence of possible recombination products.

<sup>a</sup> Four or more sequences are needed to compute Tajima's *D* and Fu and Li's statistics.

<sup>b</sup> Not significant, but  $0.1 > P > 0.05$ .

<sup>c</sup> Significant,  $P < 0.05$ .

ered undescribed species in the genus *Sphaerobolus*. Based on our morphological investigations, the possibility of this new lineage being *S. stellatus* var. *giganteus* can be ruled out. Most notably, the new lineage has much smaller fruiting bodies and gleba ( $\leq 1$  mm) than those of *S. stellatus* (1.5–2.5 mm) and *S. stellatus* var. *giganteus* (3–5 mm), and it differs in mycelial macromorphology as well, almost always producing strand-like mycelial cords on OA instead of the more variable fluffy and patchy mycelium of *S. stellatus* (Geml, unpublished data). Detailed micromorphological investigations are under way to detect characteristic differences between species.

Only three phylogenetic species, representing the three major lineages detailed above, could be detected as congruent nodes in all gene genealogies. In general, genealogies indicated little phylogeographic structure. In numerous instances, isolates from the same geographic region appeared in divergent portions of the gene tree, while isolates from geographically distant localities often clustered together. This suggests recent dispersal of isolates over great geographic distances, possibly by human activities, including the movement of livestock and wood products between distant areas. The dual coprophilous and lignicolous ecology of *Sphaerobolus* species provides many possible dispersal scenarios.

Levels of sequence divergence among *Sphaerobolus* species are similar to what is observed within morphologically diverse genera of basidiomycetes. Despite the small number of known species, values of pairwise HKY85 distances found in the first 900 bp in the 5' of the LSU of *Sphaerobolus* (5.35, 5.38, 8.12% between *S. stellatus* and *S. iowensis*, *S. iowensis* and the undescribed species, and *S. stellatus* vs. the undescribed species, respectively) are comparable to the greatest infrageneric distance values in many other homobasidiomycete genera (e.g., *Pleurotus purpureus* vs. *P. djamor* 10.62%, *Agaricus arvensis* vs. *A. bisporus* 5.72%, and *Coprinopsis lagopides* vs. *Coprinopsis friesii* 10.79%; data generated by Moncalvo et al., 2002). Despite their phylogenetic divergence, there are remarkably few obvious morphological differences between the two described *Sphaerobolus* species, while the undescribed taxon appears to have somewhat distinct macro- and micro-morphological features (Geml, unpublished data).

Some regions of the EF 1- $\alpha$  and ITS sequences were difficult to align between species, another indication of relatively deep divergence, whereas closely related species of fungi tend to have easily alignable EF 1- $\alpha$  and ITS sequences (e.g., see Geiser et al., 2001a,b; Geml and Royse, 2002; O'Donnell et al., 1998; Peintner et al., 2003). There may be a large number of unsampled taxa, extinct or extant, that would fill these phylogenetic gaps. In addition, no obvious sign of increased mutation rate was observed as unusually long branches in any of the genes examined in suprageneric phylogenetic analyses,

Table 3  
Polymorphism and divergence between species of *Sphaerobolus*

Locus/species comparison	Average no. of difference between species		Fixed differences	Polym. 1 monom. 2	Polym. 2 monom. 1	Shared polym.	Polym./divergence
	$\pi$ (per site)	$k$ (per site)					
<b>mtSSU</b>							
<i>iow-ste</i>	0.05133	0.09788	51	1	0	0	0.00633
<i>iow-und</i>	0.02399	0.08817	45	1	0	0	0.00703
<i>ste-und</i>	0.02144	0.07076	38	0	0	0	0 <sup>a</sup>
<b>ITS</b>							
<i>iow-ste</i>	0.06207	0.11715	75	7	2	0	0.03756
<i>iow-und</i>	0.04060	0.14777	88	6	10	0	0.02978
<i>ste-und</i>	0.04700	0.15252	95	1	10	0	0.00944
<b>EF 1-<math>\alpha</math></b>							
<i>iow-ste</i>	0.07932	0.13976	55	16	41	2	0.08085
<i>iow-und</i>	0.04265	0.14091	60	18	35	0	0.08019
<i>ste-und</i>	0.05566	0.12377	35	42	35	2	0.24101

The first column gives three-letter abbreviations for species compared. In the next two columns the average number of nucleotide differences are given per site for the total data ( $\pi$ ) and between species (divergence,  $k$ ). Subsequently, four columns indicate the number of fixed differences between species, sites that are polymorphic in the first species (appearing on the left of the comparison column) but monomorphic in the second (on the right), sites that are polymorphic in the second species but monomorphic in the first, and sites that are polymorphic in both species, respectively. The right column indicates the ratio of polymorphism to divergence ( $\theta/k$ ) where polymorphism data are from the first species.

<sup>a</sup> There was no polymorphic site in the mtSSU region in *S. stellatus*, therefore  $\theta = 0$ .

confirming the ancient split between the lineages as the most likely explanation for the divergences observed between them.

Suprageneric phylogenetic analyses of the mtSSU and LSU datasets containing representatives of related genera of the gomphoid–phalloid clade of Homobasidiomycetes indicated that the genus *Sphaerobolus* is monophyletic, and suggested that the undescribed taxon likely is more closely related to *S. stellatus* than to *S. iowensis*, despite the fact that nucleotide differences between *S. iowensis* and the undescribed species were sometimes smaller than those between *S. stellatus* and the undescribed taxon. Increasing the number of sampled and analyzed isolates in all species should clarify this question.

In this paper, we revealed the existence of at least three deeply divergent phylogenetic species in the genus *Sphaerobolus* as opposed to the only two widely accepted species providing a phylogenetic framework for the genus that can serve as base for future research on *Sphaerobolus*. Since our analyses included isolates predominantly from North America, samples from other areas are needed to further elucidate the genetic diversity and phylogenetic structure of the genus. Among these, tropical regions are of particular interest because of their high biodiversity and largely unknown fungal flora. In our future analyses, we are going to increase the number of isolates sampled to learn more about the evolution and population structure of *Sphaerobolus* species. In addition, we will use the revealed phylogenetic relationships to observe a posteriori variation in morphological and/or biological characteristics corresponding to the separate lineages that otherwise would be much more difficult to

detect. This approach was used earlier to reveal genetic and later morphological variation in cryptic species, such as *Aspergillus flavus* (Geiser et al., 2001a,b).

### Acknowledgments

This research was supported by the Pennsylvania Department of Agriculture and the Deep Hypha Project (NSF 0090301). The authors thank the following persons for providing cultures: Elizabeth Brantley, Richard Hanlin, David Hibbett, David McLaughlin, Lori Trummer, and Kazumasa Yokoyama. Many thanks go to Frank Kauff, Jolanta Miadlikowska, and François Lutzoni for generously providing the ARC software. The authors also thank the two anonymous reviewers for their helpful comments.

### References

- Alasoadura, S.O., 1963. Fruiting in *Sphaerobolus* with special reference to light. *Ann. Bot. N.S.* 27, 123–145.
- Baayen, R.P., O'Donnell, K., Breeuwsma, S., Geiser, D.M., Waalwijk, C., 2001. Molecular relationships of fungi within the *Fusarium redolens*–*F. hostae* clade. *Phytopathology* 91 (11), 1037–1044.
- Brantley, E.A., Davis, D.D., Kuhns, L.J., 2001a. Biological control of the artillery fungus, *Sphaerobolus stellatus*, with *Trichoderma harzianum* and *Bacillus subtilis*. *J. Environ. Hortic.* 19, 21–23.
- Brantley, E.A., Davis, D.D., Kuhns, L.J., 2001b. Influence of mulch characteristics on the sporulation of the artillery fungus *Sphaerobolus stellatus*. *J. Environ. Hortic.* 19, 89–95.
- Bruns, T.D., White, T.J., Taylor, J.W., 1991. Fungal molecular systematics. *Annu. Rev. Ecol. Syst.* 22, 525–564.
- Buller, A.H.R., 1933. *Researches on Fungi*, Vol. 5. Longmans, Green & Co., London.

- Dykstra, M.J., 1982. A cytological examination of *Sphaerobolus stellatus* fruiting bodies. *Mycologia* 74 (1), 44–53.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Fletcher, M., Cooke, C., 1984. Carbohydrate changes in the developing sporophore of *Sphaerobolus stellatus*. *Trans. Br. Mycol. Soc.* 82 (2), 366–369.
- Fu, Y.-X., Li, W.-H., 1993. Maximum likelihood estimation of population parameters. *Genetics* 134, 1261–1270.
- Geiser, D.M., Dorner, J.W., Horn, J.W., Taylor, J.W., 2001a. The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*. *Fung. Gen. Biol.* 31, 169–179.
- Geiser, D.M., Juba, J.H., Wang, B., Jeffers, S.N., 2001b. *Fusarium hostae* sp. nov., a relative of *F. redolens* with a *Gibberella* teleomorph. *Mycologia* 93, 670–678.
- Geml, J., Royse, D.J., 2002. Molecular phylogeny and cultivation of *Agaricus* species. In: Proceedings of the IV International Conference on Mushroom Biology and Mushroom Products, February 20–23, 2002, Cuernavaca, Mexico, pp. 111–120.
- Hasegawa, M., Kishino, H., Yano, T.A., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 21, 160–174.
- Hawksworth, D.L., Kirk, P.M., Sutton, B.C., Pegler, D.N., 1996. *Ainsworth & Bisby's Dictionary of the Fungi*, eighth ed. Oxford University Press, NY.
- Huelsenbeck, J.P., Ronquist, F., 2001. Mr.Bayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Hibbett, D.S., Pine, E.M., Langer, E., Langer, G., Donoghue, M.J., 1997. Evolution of gilled mushrooms and puffballs inferred from ribosomal DNA sequences. *Proc. Natl. Acad. Sci. USA* 94, 12002–12006.
- Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA—molecular evolution and phylogenetic inference. *Q. Rev. Biol.* 66, 410–453.
- Hopple, J.S., Vilgalys, R., 1999. Phylogenetic relationships in the mushroom genus *Coprinus* and dark-spored allies based on sequence data from the nuclear gene coding for the large ribosomal subunit RNA: divergent domains, outgroups, and monophyly. *Mol. Phylogenet. Evol.* 13, 1–19.
- Ingold, C.T., 1972. *Sphaerobolus*: the story of a fungus. *Trans. Br. Mycol. Soc.* 58, 179–195.
- Kimura, M., 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, MA.
- Kirk, P.M., Cannon, P.F., David, J.C., Stalpers, J.A., 2001. *Dictionary of the Fungi*, ninth ed. CAB International, Wallingford, UK.
- Lehman, R.D., 1985. Black spots on houses—an insect or disease problem. *Penna. Dept. Agric. Bur. Pl. Ind. Reg. Hortic.* 11, 15–16.
- Lutzoni, F., Vilgalys, R., 1995. Integration of morphological and molecular data sets in estimating fungal phylogenies. *Can. J. Bot. Suppl.* 73, S649–S659.
- Lutzoni, F., Wagner, W., Reeb, V., Zoller, S., 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. *Syst. Biol.* 49, 628–651.
- Miadlikowska, J., Lutzoni, F., Goward, T., Zoller, S., Posada, D., 2003. New approach to an old problem: Incorporating signal from gap-rich regions of ITS and nrDNA large subunit into phylogenetic analyses to resolve the *Peltigera canina* species complex. *Mycologia* 95, 1181–1203.
- Micheli, P.A., 1729. *Nova plantarum genera iuxta Tournefortii methodum disposita*. Florence, Italy (reprinted 1976, Richmond Publishing, Richmond).
- Moncalvo, J.-M., Vilgalys, R., Redhead, S.A., Johnson, J.E., James, T.Y., Aime, M.C., Hofstetter, V., Verduin, S.J.W., Larsson, E., Baroni, T.J., Thorn, R.G., Jacobsson, S., Cléménçon, H., Miller Jr., O.K., 2002. One hundred and seventeen clades of euagarics. *Mol. Phylogenet. Evol.* 23, 357–400.
- Moncalvo, J.M., Wang, H.H., Hseu, R.S., 1995. Phylogenetic relationships in *Ganoderma* inferred from the internal transcribed spacers and 25S ribosomal DNA sequences. *Mycologia* 87, 223–238.
- Nei, M., Li, W.-H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76, 5269–5273.
- O'Donnell, K., 1996. Progress towards a phylogenetic classification of *Fusarium*. *Sydowia* 48 (1), 57–70.
- O'Donnell, K., Kistler, H.C., Cigelnik, E., Ploetz, R.C., 1998. Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Proc. Natl. Acad. Sci. USA* 95, 2044–2049.
- O'Donnell, K., Lutzoni, F.M., Ward, T.J., Benny, G.L., 2001. Evolutionary relationships among mucoralean fungi (Zygomycota): evidence for family polyphyly on a large scale. *Mycologia* 93 (2), 286–296.
- Peintner, U., Moser, M.M., Thomas, K.A., Manimohan, P., 2003. First records of ectomycorrhizal *Cortinarius* species (Agaricales, Basidiomycetes) from tropical India and their phylogenetic position based on rDNA ITS sequences. *Mycol. Res.* 107, 485–494.
- Pine, E.M., Hibbett, D.S., Donoghue, M.J., 1999. Phylogenetic relationships of cantharelloid and clavarioid Homobasidiomycetes based on mitochondrial and nuclear rDNA sequences. *Mycologia* 91, 944–963.
- Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14 (9), 817–818.
- Rozas, J., Rozas, R., 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15, 174–175.
- Swofford, D.L., 2002. *PAUP\*. Phylogenetic Analysis Using Parsimony (\* and other methods)*, 4.0b4a. Sinauer Associates, Sunderland, MA.
- Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- 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.
- Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geiser, D.M., Hibbett, D.S., Fisher, M.C., 2000. Phylogenetic species recognition and species concepts in fungi. *Fung. Gen. Biol.* 31, 21–32.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25 (24), 4876–4882.
- Ulloa, M., Hanlin, R.T., 2000. *Illustrated Dictionary of Mycology*. APS Press, St. Paul, MN.
- Walker, L.B., 1927. Development and mechanism of discharge in *Sphaerobolus iowensis* and *S. stellatus* Tode. *J. Eli. Mitch. Sci. Soc.* 42, 151–178.
- Walker, L.B., Anderson, E.N., 1925. Relation of glycogen to spore-ejection. *Mycologia* 15, 154–159.
- Watterson, G.A., 1975. On the number of segregating sites in genetic models without recombination. *Theor. Popul. Biol.* 10, 256–276.
- White, T.J., Bruns, T., Lee, S., Taylor, J.W., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), *PCR Protocols: A Guide to Methods and Applications*. Academic Press, New York, pp. 315–322.