C.1. RESULTS FROM PREVIOUS NSF SUPPORT.

The PIs of the proposed research have each received NSF support within the last five years. One PI on this project (JW Spatafora) is also a Co-PI on the ongoing "Deep Hypha" Research Coordination Network (RCN) (http://ocid.NACSE.ORG/research/deephyphae/) (DEB-0090301, Co-PIs, J. Taylor and M. Blackwell, 2001-2005), and another PI (D. Hibbett) is a member of the Deep Hypha Steering Committee. Collectively, support to the PIs of the present proposal in the last five years has resulted in 110 unique peer-reviewed publications and numerous presentations, and has supported 23 graduate students, 13 post-doctoral fellows, 21 undergraduates, and 5 high school students. Because of space limitations, only basic information about one research grant to each PI (all from the Systematics and Population Biology Program), and selected publications, are presented below.

C.1.1 David Hibbett: Morphological and ecological diversification in the homobasidiomycetes: a molecular phylogenetic analysis. (DSH, PI) DEB-9903835 (Sept., 1999-Aug., 2002) $190,000.

Selected publications (since 1997, out of 18):

Hibbett, D. S., L.-B. Gilbert, and M. J. Donoghue. 2000. Evolutionary instability of ectomycorrhizal symbioses in basidiomycetes. *Nature* 407: 506-508.

Hibbett, D. S., and M. J. Donoghue. 2001. Analysis of correlations among wood decay mechanisms, mating systems, and substrate ranges in homobasidiomycetes. *Syst. Biol.* 50: 215-242.

Hibbett, D. S., and R. G. Thorn. 2001. Basidiomycota: Homobasidiomycetes. Pp. 121-168 *in*: The Mycota, vol. VII part B, Systematics and Evolution (D. J. McLaughlin, E. G. McLaughlin, and P. A. Lemke, eds.). Springer Verlag.

Binder, M., and D. S. Hibbett. 2002. Higher level phylogenetic relationships of homobasidiomycetes inferred from four rDNA regions. *Molecular Phylogenetics and Evolution* 22:76-90.

Taylor, J. W., J. Spatafora, K. O'Donnell, F. Lutzoni, T. James, D. S. Hibbett, D. Geiser, T. D. Bruns, and M. Blackwell. The Kingdom Fungi. *In:* Companion volume to symposium on "Assembling the Tree of Life: Science, Relevance and Challenges", American Museum of Natural History May 30-June 1, 2002 (J. Cracraft and M. J. Donoghue, eds.). In preparation.

C.1.2 François Lutzoni: CAREER: Using multigene phylogenies to solve early ascomycete relationships and reconstruct the origin and losses of the lichen symbiosis. (FL, PI) DEB-0133891 (Feb., 2002-Jan. 2007) $680,000.

Selected publications (since 1997, out of 24)

Lutzoni, F., P. Wagner, V. Reeb, and S. Zoller. 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. *Syst. Biol.* 49:628-651.

Lutzoni, F., M. Pagel, V. Reeb. 2001. Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* 411:937-940.

Kauff, F. and F. Lutzoni. 2002. Phylogeny of the Gyalectales and Ostropales (Ascomycota, Fungi): among and within order relationships based on nuclear ribosomal RNA small and large subunits. *Molecular Phylogenetics and Evolution* (in press).

Barker, F. K., and F. Lutzoni. 2002. The utility of the incongruence length difference test. *Syst. Biol.* 51(3): In press.

Alfaro, M, S. Zoller, F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* (in review).

C.1.3 David McLaughlin: Molecular systematics and evolution of simple-septate basidiomycetes. (DJM, PI, E. Swann, Co-PI) DEB 9318232 (Feb., 1994-June,1999) $180,000.

Selected publications (since 1995, out of 11)

McLaughlin, D.J., E.M. Frieders and H. Lü. 1995. A microsopist's view of heterobasidiomycete phylogeny. Studies in Mycology 38:91-109.

Frieders, E.M., and D.J. McLaughlin. 1996. Mitosis in the yeast phase of *Agaricostilbum pulcherrimum* and its evolutionary significance. Can. J. Bot. 74:1392-1406.

Swann, E.C., E.M. Frieders, and D. J. McLaughlin. 1999. *Microbotryum*, *Kriegeria*, and the changing paradigm in basidiomycete classification. Mycologia 91: 51-66.

Swann, E.C., E.M. Frieders, and D. J. McLaughlin. 2001. Urediniomycetes. In McLaughlin, D.J., E.G. McLaughlin, and P.A. Lemke (eds.). The Mycota. VIIB. Systematics and Evolution. Springer Verlag. Pp. 37-56.

Frieders, E.M., and D.J. McLaughlin. 2001. The heterobasidiomycete moss parasites *Jola* and *Eocronartium* in culture: cytology, ultrastructure, and anamorph. Mycol. Res. 105: 734-744.

C.1.4 Joseph Spatafora: A five gene region phylogeny for Cordyceps and the Clavicipitaceae (Ascomycota). (JWS, PI) DEB-0129212 (Mar., 2002-Feb., 2005) $260,000.

Selected publications (since 1997, out of 20):

Suh, S.-O., J. W. Spatafora, G.R.S. Ochiel, H. C. Evans, and M. Blackwell. 1998. Molecular phylogenetic study of a termite pathogen, *Cordycepioideus bisporus*. *Mycologia* 90: 611-617.

Sullivan, R., M.S. Bergen, R. Patel, G.F. Bills, S.C. Alderman, J.W. Spatafora, and J.F. White, Jr. 2001. *Neoclaviceps monostipa*: features of an enigmatic clavicipitalean fungus and the phyletic status of the anamorphic genus *Ephelis*. *Mycologia* 93: 90-99.

Sung, G.-H., J.W. Spatafora, R. Zare, K.T. Hodge, and W. Gams. 2001. A revision of *Verticillium* sect. *Prostrata.* II. Phylogenetic analyses of SSU and LSU nuclear rDNA sequences from anamorphs and teleomorphs of the Clavicipitaceae. *Nova Hedwigia* 72: 29-46.

Spatafora, J. W. In press. Evolution of Ascomycota-Arthropoda Symbioses. *In*: Symbiosis. Ed. J. Seckbach. Kluwer, Dordrecht.

Taylor, J. W., J. Spatafora, K. O'Donnell, F. Lutzoni, T. James, D. S. Hibbett, D. Geiser, T. D. Bruns, and M. Blackwell. The Kingdom Fungi. *In:* Companion volume to symposium on "Assembling the Tree of Life: Science, Relevance and Challenges", American Museum of Natural History May 30-June 1, 2002 (J. Cracraft and M. J. Donoghue, eds.). In preparation.

C.1.5 Rytas Vilgalys: Phylogenetic systematics and diversity of the Agaricales. (RV, PI) DEB-0076023 (Aug., 2000-Jul., 2002) $225,000.

Selected publications (since 1997, out of 37):

James, T. Y., D. Porter, C. A. Leander, R. Vilgalys, and J. E. Longcore 2000. Molecular phylogenetics of the Chytridiomycota support the utility of subcellular data in chytrid systematics. Can. J. Bot. 78: 336-350

Moncalvo, J-M., Lutzoni, F. M., Rehner, S. A., Johnson, J., and R. Vilgalys. 2000. Phylogenetic relationships of agaric Fungi based on large ribosomal subunit DNA sequences. *Syst. Biol.* 49: 278-305

J.-M. Moncalvo, R. Vilgalys, S.A.Redhead, J.E.Johnson, T.Y.James, M.C. Aime, V. Hofstetter, S.J. W.Verduin, E. Larsson, T.J.Baroni, R.G. Thorn, S. Jacobsson, H. Clémençon, and O.K. Miller, Jr. 2002. One Hundred and Seventeen Clades of Euagarics. *Mol. Phyl. Evol*. 19 (in press).

Gonzalez, D., D. E. Carling, S. Kuninaga, R. Vilgalys, M. A. Cubeta. 2001. Ribosomal DNA systematics of *Ceratobasidium* and *Thanatephorus* with *Rhizoctonia* anamorphs *Mycologia* 93: 1138-1150

Peintner, U., N. L. Bougher, M. A. Castellano, J.-M. Moncalvo, M. M. Moser, J. M. Trappe, and R. Vilgalys. 2001. Multiple origins of sequestrate Fungi related to *Cortinarius* (Cortinariaceae). *American Journal of Botany* 88: 2168-2179

C.2 INTRODUCTION

Fungi make up one of the major clades of Life. There are roughly 80,000 described species of Fungi, but the actual number of species has been estimated at approximately 1.5 million (Hawksworth 1991, 2001; Hawksworth et al. 1995). If this estimate is correct, then less than 6% of all Fungi have been documented. Fungi play pivotal ecological roles: As **saprotrophs**, they are important in the cycling of nutrients, especially the carbon that is sequestered in wood and other plant tissues. As **pathogens** and **parasites**, they attack virtually all groups of organisms, including plants, other Fungi, and animals, including humans. As mutualistic **symbionts** (e.g., lichens, mycorrhizae), Fungi have enabled diverse organisms to exploit novel habitats and resources. Fungi have been found in every ecosystem where they have been sought, including deserts, glacial ice, and deep-sea thermal vent communities (Jacobson et al. 1999; DePriest et al. 2000; Kohlmeyer & Kohlmeyer 1979; van Dover & Lichtwardt 1986). For all these

reasons, a robust fungal phylogeny is necessary for understanding the history of Life, including the evolution of ecosystems. A fungal phylogeny will also provide resources for applied purposes, including plant pathology, medical mycology, and drug discovery, and it will facilitate the discovery of the many fungal species that remain undocumented.

In recent years, mycologists have made major advances in resolving fungal phylogeny. The advances have not been uniform, however. Most researchers have concentrated on the Ascomycota and Basidiomycota, which contain approximately 95% of the known species, but relatively few workers have addressed relationships of the Chytridiomycota and Zygomycota (O'Donnell et al. 2001), both of which appear to be non-monophyletic (James et al. 2000; Tanabe et al. 2000). In addition, lack of coordination among research groups has often resulted in the creation of non-overlapping datasets, which are informative for their focal group but are not easily combined. As the number of genes used in fungal phylogenetics increases, the challenge of integrating datasets will only grow more difficult.

We propose a collaborative project among five laboratories, titled "Assembling the Fungal Tree of Life" (AFTOL), which will develop a higher-level phylogenetic framework for Fungi, using molecular and non-molecular characters. Four laboratories will coordinate sequencing and analysis of seven loci from approximately 1500 representative species from all major groups of Fungi, and one laboratory will generate non-molecular characters that have shown promise for higher-level fungal phylogenetics. Web-accessible databases for molecular and non-molecular characters in fungal phylogenetics will also be developed. The proposed research will require a high degree of community participation. This will be facilitated by the Deep Hypha RCN, which provides a forum for the fungal systematics community through two annual meetings. Every effort will be made to make the proposed research as open and inclusive as possible. All the data produced in this project will be released via the world-wide web as soon as they have been generated, and participants who donate materials will be invited to co-author publications reporting new data derived from their contributions.

Training and outreach activities will take several forms: 1) each of the participating laboratories will support one resident graduate student and one post-doctoral fellow; 2) each lab will also host visiting graduate students (rarely, post-docs), including foreign students, who will be supported for 3-12 month training and research visits; 3) REU supplements will be sought to support summer research opportunities for undergraduates; 4) each lab will host a workshop every summer for local high school science educators, who will receive training in basic mycology as well as recent advances in fungal phylogeny; 5) a "Fungal K-12" web site will be developed with information and resources that will enable K-12 educators to bring Fungi into their curricula.

C.3 BACKGROUND INFORMATION

C.3.1 History and prospects of fungal molecular systematics. The science of fungal systematics has been utterly transformed by the application of molecular techniques. Early work in fungal molecular systematics (reviewed by Bruns et al. 1991; and Hibbett 1992) focused largely on ribosomal RNAs (rRNA) and the genes that encode them (rDNA). These were attractive because the rDNA occurs in high copy number, facilitating cloning and Southern hybridization studies, and the rRNA can be sequenced directly with reverse transcriptase and other methods (e.g., Anderson et al. 1987; Blanz & Gottschalk 1984; Kohn et al. 1988; Guadet et al. 1989). The development of PCR in the late 1980s (Saiki et al. 1998), and, later, automated methods of DNA sequencing, ushered in a flood of new data for fungal phylogenetics. Concurrent advances in computer science and methods of phylogenetic analysis were equally important in the growth of the discipline. White et al. (1990) published a seminal paper listing primers for PCR amplification and sequencing of fungal nuclear and mitochondrial rDNA, and since then these regions have been by far the most frequently studied in fungal molecular systematics. Indeed, most of what is now understood about fungal phylogeny is based on rDNA, especially nuclear rDNA. In addition, rDNA internal transcribed spacer sequences (ITS) and other rDNA regions are being used with increasing frequency in environmental studies of fungal diversity (e.g., Bruns et al 1998; Horton et al 1999; Vandenkoornhuyse et al. 2002; O'Brien et al. 2002). It is impossible in this short space to provide a

complete review of the progress in fungal molecular systematics that has occurred in the last fourteen years. Much of this work is summarized by Taylor et al. (in press) and in the recent *Mycota* vol. VII, edited by McLaughlin et al. (2001).

To achieve a comprehensive phylogeny for the Fungi, and to develop tools for the discovery and detection of Fungi in nature, it is imperative that the fungal rDNA database continues to grow. Nevertheless, it has become clear that rDNA will not solve all the problems of fungal phylogeny. For example, Binder & Hibbett (2002) analyzed four regions of mitochondrial and nuclear rDNA in the homobasidiomycetes, but found that some basal nodes remained poorly resolved. Evolutionary rate heterogeneity and the occasional existence of paralogous copies of rDNA (O'Donnell et al. 1998) have also been cited as deficiencies of rDNA for fungal phylogenetics. Finally, there is a growing realization that comparison of multiple gene phylogenies is a powerful method for assessing species limits in Fungi (Taylor et al. 2000). Consequently, many fungal systematists have begun to explore various protein-coding genes, including glyceraldehyde 3-phosphate dehydrogenase (Berbee et al. 1999), β-tubulin (β-tub, Landvik et al 2001; Leuchtmann & Schardl 1998; O'Donnell et al. 1998; Schardl et al. 1994), elongation factor 1-α (ef 1-α, O'Donnell et al 2001), RNA polymerase I subunit B (RPB1, Matheny et al. 2002), RNA polymerase II subunit B (RPB2, Liu et al. 1999), and mitochondrial ATP6 (Kretzer et al. 1999). The number of loci that are used in fungal molecular systematics will continue to grow, especially as more complete fungal genome sequences are produced (Bennett & Arnold 2001).

What is ultimately needed is to combine information from more than one locus into one dataset, an approach that had considerable success, especially in green plant phylogenetics (e.g., Pryer et al 2001; Qiu et al. 1999; Soltis et al. 1999) and kingdom-level studies of eukaryotes (Baldauf & Palmer 1993; Baldauf et al 2000). However, obtaining datasets from different genes for the same set of species is very challenging, and the number of higher-level fungal systematics studies that have utilized two or more loci or data partitions (e.g., Bhattacharya et al. 2000; Binder & Hibbett 2002; Forget et al 2002; Hibbett et al 1997; Kauff & Lutzoni 2002; Lumbsch et al 2002; Lutzoni et al. 2001; O'Donnell et al 2001; Paquin et al 1997; Spatafora et al 1998; Suh & Blackwell 1999; Sung et al 2001) remains in the minority of studies. For this approach to succeed, coordination among members of the fungal systematics community will be increasingly important. Access to cultures, specimens, and especially DNA samples will be necessary to insure that the same individuals can be studied for multiple loci.

C.3.2. Molecular perspectives on monophyly and higher-level relationships of Fungi. Molecular phylogenetic studies have greatly refined our understanding of the limits of the Fungi. Several groups that have traditionally been included in the Fungi have been shown to belong to other clades, including the oomycetes, labyrinthulomycetes, and plasmodiophoromycetes (all of which have been shown to be in the Heterokonta, also called Stramenopila), dictyostelid and plasmodial slime molds (Mycetozoa), and the Amoebidiales, which is a group of arthropod symbionts that has been shown to be related to certain protists (Baldauf & Doolittle 1997; Benny & O'Donnell 2000; Leipe et al. 1994; Saunders et al. 1997). Molecular analyses have also shown that certain taxa that have not previously been placed in the Fungi are actually members of this clade, including *Pneumocystis carinii*, which is a devastating pathogen of AIDS patients (Edman et al. 1988), and the microsporidia, which are unusual intracellular parasites that lack mitochondria (Baldauf et al. 2000; Barr 1992; Cavalier-Smith et al. 1994; Keeling et al. 2000). Analyses of rDNA and various protein-coding genes suggest that the Fungi are more closely related to Animalia, choanoflagellates, and Mycetozoa than they are to other groups of eukaryotes (Baldauf et al. 2000; Berbee & Taylor 2001; Heckman et al. 2001; Wainwright et al. 1993). The precise branching order among these groups is not well resolved, however.

C.3.3. Molecular perspectives on phylogenetic relationships within the Fungi. Historically, the Fungi have been divided into four phyla: Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (reviewed in Alexopoulos et al 1996). The Chytridiomycota includes the only flagellated members of the Fungi, many of which are unicellular and aquatic. The remaining phyla are mostly terrestrial and filamentous (although yeasts and aquatic forms also occur). Earlier taxonomic systems delimited the Chytridiomycota and Zygomycota as separate groups, based principally on the presence or absence of

flagellated cells, respectively. However, molecular characters, primarily ssu-rDNA sequences, suggest that neither group is monophyletic (Sugiyama 1998, Tanabe et al. 2000, James et al. 2000). Together the Chytridiomycota and Zygomycota form a poorly resolved, paraphyletic group at the base of the Fungi that consists of possibly up to 10 separate lineages including 1) **Glomales,** 2) **Endogonales/ Mortierellaceae,** 3) **the core Mucorales,** 4) **Entomophthorales,** 5) **Zoopagales/Kickxellales/Harpellales,** 6) **Dimargaritales**, 7) **core chytrid clade,** 8) **Monoblepharidales,** 9) **Blastocladiales, and** 10) **Eccrinales/ Asellariales** (Fig. 1; Bruns et al. 1992; James et al. 2000; Jensen et al. 1998; Nagahama et al. 1995; O'Donnell et al. 2001; Tanabe et al. 2000). The lack of resolution of the Chytridiomycota and Zygomycota makes it difficult to resolve some of the key events in fungal evolution, such as the transition(s) to land, loss(es) of flagella, and gain(s) of the filamentous habit.

Phylogenetic studies of nuclear rDNA suggest that the Ascomycota and Basidiomycota are both monophyletic and are sister taxa (Bruns et al. 1992). Together, they comprise a terminal clade that has been classified as the Dikaryomycotina, and which is informally referred to as the "higher fungi". The Basidiomycota includes three classes: **Urediniomycetes** (rusts and relatives), **Ustilaginomycetes** (smuts), and **Hymenomycetes** (fleshy Fungi). Monophyly of these groups is supported by rDNA studies, but the interrelationships of the three clades is equivocal (Swann & Taylor 1993, 1995; Swann et al. 1999, 2001). The major clades within the Urediniomycetes, Ustilaginomycetes, and Hymenomycetes have largely been delimited through analyses of rDNA (Fig.1; Bauer et al. 2001; Hibbett & Thorn 2001; Swann et al. 2001), but relationships among and within these clades are often poorly resolved. Sampling has been particularly intensive in the Hymenomycetes (Hibbett et al 2000; Langer in press; Boidin et al. 1998), especially the euagarics clade of the homobasidiomycetes (Moncalvo et al. 2000, 2002).

Comparison of the genes that encode for the nuclear ribosomal RNAs and the gene family of RNA polymerase, especially RPB*2*, suggest that the Ascomycota possesses three major subgroups (Fig 1; Berbee & Taylor 1993; Bruns et al. 1992; Gargas et al. 1995; Gernandt et al 2001; Spatafora 1995; Liu et al. 1999; Lutzoni et al. 2001). In the most recent classification (Eriksson et al. 2001) the three groups are the Subphylum **Taphrinomycotina** (= Class Archiascomycetes), the Subphylum **Saccharomycotina** (= Class Hemiascomycetes), and the Subphylum **Pezizomycotina** (= Class Euascomycetes). The Taphrinomycotina is supported as the most basal lineage of the Ascomycota, although its monophyly is not supported in all analyses suggesting it may comprise a paraphyletic assemblage of early diverging lineages of the phylum. The Saccharomycotina comprises the organisms most biologists recognize as yeasts or "true yeasts" and is home to one of the best known species of Fungi, *Saccharomyces cerevisiae*, better known as the baker's yeast (and the first eukaryote to have its complete genome sequenced). Although most Saccharomycotina are primarily unicellular, numerous species do make abundant hyphae, but none produce ascomata (Barnett et al. 1990). The Pezizomycotina contains the majority of lichenized and nonlichenized species of Ascomycota, most of which produce ascomata. Although it contains the majority of described fungal species, the relationships of the major clades of the Pezizomycotina remain poorly resolved (Gernandt et al 2001, Lutzoni et al 2001). A preliminary study by Lutzoni et al. (in prep.) based on a combined analysis of nuc SSU, nuc LSU rDNA and RPB2 for the same set of 60 species, representing most major lineages within the Ascomycota, demonstrate that adding RPB2 to the two rRNA genes greatly reduces phylogenetic uncertainty near the base of the Pezizomycotina, but is insufficient to completely resolve this problematic region of the tree of the Ascomycota. It is also expected that as more taxa are added, support values will go down for many of these short internodes. This, and other similar cases throughout Fungi, justify our proposal to sequence three additional loci, RPB1, EF-1α, and ATP6.

C.3.4 Nonmolecular characters in higher-level fungal phylogeny. Subcellular characters (e.g., those of flagella and septal pores; Beckett et al. 1974) and biochemical characters (e.g., biosynthetic pathways and cell wall composition; Bartnicki-Garcia, 1970, 1987) provided insight into phylogenetic relationships of Fungi long before the advent of molecular characters (Alexopoulos et al. 1996). For example, in the Chytridiomycota and Zygomycota, variation in motile cell structure, spindle pole body (SPB), and nuclear division features suggested that these groups are not monophyletic (Barr 2001; Heath 1986), which has been confirmed with molecular analyses (James et al. 2000; Nagahama et al. 1995). In the

Ascomycota and Basidiomycota, meiosporangia, their associated spore discharge mechanisms, phylum-specific roles of dikaryotic hyphae, and septum organization suggested the existence of clades that have also been supported by molecular analyses (Alexopoulos et al. 1996; Bruns et al. 1992). Within each phylum a wide array of subcellular characters appear to be phylogenetically informative at class and lower taxon levels (Pfister & Kimbrough 2001; Swann et al. 1999).

Despite their apparent informativeness, there have been major problems with the few higher-level phylogenetic analyses of subcellular characters in Fungi. For example, previous analyses of SPB and nuclear characters were performed without reference to other phylogenetically informative features (Heath 1986). The lack of data for many taxa is a particularly serious limitation (McLaughlin et al. 1995a). Because of a lack of adequate sampling, some workers have scored certain structural and biochemical characters that have been observed in only one or a few species for entire families and orders, which have later been shown not to be monophyletic (Tehler 1988). To take full advantage of subcellular and biochemical characters for resolving fungal phylogeny, it will be necessary to compile and reevaluate published characters, as well as generate new data. In some cases, reanalyses of previously reported characters using improved preservation methods will be needed. Significant gaps exist for many characters in all phyla (e.g, for SPB and nuclear division characters), but especially in Zygomycota and Ascomycota. These gaps in the data become more apparent as our understanding of monophyletic groups is clarified by molecular analysis.

C.3.5 Summary—the state of fungal phylogenetics. Considerable progress has been made in reconstructing fungal phylogeny through studies of sub-cellular characters and molecular characters, specifically rDNA sequences. Nevertheless, fewer than 900 species (~1% of described species) of Fungi have been included together in any single phylogenetic reconstruction using rDNA, and there are significant “holes” in the subcellular data (Moncalvo et al. 2002; Tehler et al 2000). To approach a complete hypothesis of fungal phylogeny, it will be necessary to expand the rDNA database and fill in the missing data for sub-cellular characters. It will also be necessary to generate comprehensive datasets for protein-coding loci, and to do so in a way that makes the new datasets combinable with each other and with the rDNA data.

Two of the major obstacles that have prevented a comprehensive survey of the diversity of Fungi include accessibility of specimens and the sheer amount of phylogenetic diversity relative to the number of fungal systematists. Many groups of Fungi are best known by the taxonomists and ecologists who study them, but whose primary research interests do not include molecular systematics. The proposed research will advance a world-wide model of communication and collaboration among mycologists, facilitated by the Deep Hypha research coordination network, and it will generate comprehensive datasets of multiple genes and key subcellular characters.

C.4 RESEARCH PLAN

C.4.1 Overview of project organization

C.4.1.1 Main centers and their responsibilities. The proposed research involves the coordinated efforts of five laboratories at four universities (two public and two private), as well as participation by members of the international fungal systematics community. A timetable is provided in the Management Plan included in Supplementary Documents. The unique responsibilities of the coordinating laboratories, including database and outreach activities, are outlined below:

- Hibbett (Clark University): Collection of molecular data from Basidiomycota; creation and maintenance of a "Fungal K-12" web site for K-12 educators.
- Lutzoni (Duke University): Collection of molecular data from lichen-forming Ascomycota and allied nonlichenized taxa within the Pezizomycotina; creation and maintenance of a web-accessible database of molecular characters from all groups of Fungi.
- McLaughlin (University of Minnesota): Collection of selected non-molecular characters from all groups of Fungi; creation and maintenance of a web-accessible database of non-molecular characters from all groups of Fungi.

- Spatafora (Oregon State University): Collection of molecular data from Ascomycota: Taphrinomycotina, Saccharomycotina and nonlichenized Pezizomycotina; coordination of AFTOL and Deep Hypha, including integration of web sites.
- Vilgalys (Duke University): Collection of molecular data from Chytridiomycota and Zygomycota; oversight of DNA sequencing.

The four laboratories that are focused on molecular data will all sequence the same core set of seven genes, including nuc-lsu rDNA, nuc-ssu rDNA, ITS, RPB1, RPB2, EF-1α, and ATP6. Each of the four molecular laboratories will perform all stages of molecular phylogenetic analysis. In addition, the Vilgalys laboratory will provide a central sequencing facility (organized around an ABI 3700 capillary automated DNA sequencer) that will expedite sequencing. Non-molecular data, including subcellular and biochemical characters, will be compiled for approximately 400 species. Both molecular and non-molecular data will be disseminated via web-accessible databases, and will be released as soon as they have been generated. Additional information regarding project organization are presented in this section (Research Plan) and in the Management Plan (supplementary documentation).

C.4.1.2 Involvement of postdoctoral fellows, graduate students, and undergraduates. Each of the five participating laboratories will provide support for one post-doctoral fellow and one graduate student, who will share responsibility for collecting molecular and non-molecular data. In addition, the Lutzoni laboratory will provide support for a sixth post-doctoral fellow, who will be responsible for maintaining the fungal molecular phylogenetic database. Each of the participating laboratories will apply for REU supplements to support undergraduates, who will receive training in techniques of fungal systematics.

Support will also be provided for visiting graduate students, including international students, who are not based in any of the PI's institutions. The goals of this aspect of the AFTOL project are 1) to provide training in techniques of fungal systematics; 2) to promote cooperation among independent laboratories; and, 3) to contribute to the data collection goals of the AFTOL project. Each of the five participating laboratories will host one visiting student at a time, for periods of 3 months to one year. Thus, over the four-year period of the proposed research, 20-80 visiting students will be supported. In rare cases, visiting post-doctoral fellows will be supported, but only when there is a clear opportunity to provide training in techniques with which the post-doctoral fellow has had little experience.

C.4.1.3 Involvement of collaborators. Because the AFTOL project will involve all major groups of Fungi, it will be essential to collaborate with taxonomic specialists, who will advise regarding sampling and assist in the procurement of specimens and data. Every effort will be made to involve collaborators as full partners in the proposed research. Already, 113 collaborators, in 23 countries, have indicated that they would like to participate in the AFTOL project (see letters of support).

C.4.2 Taxon sampling.

C.4.2.1. Overview of taxon sampling. Approximately 1500 species will be sampled, including 800 Ascomycota, 400 Basidiomycota, 100 Chytridiomycota, and 200 Zygomycota. Under this sampling design, the Chytridiomycota and Zygomycota, which collectively include <5% of the described species of Fungi (Hawksworth et al. 1995), will be overrepresented. The Chytridiomycota and Zygomycota include the most basal clades of Fungi, however, and both have been shown to be nonmonophyletic (Bruns et al. 1992, James et al. 2000, Tanabe et al. 2000). Resolution of the relationships of Chytridiomycota and Zygomycota is critical to understanding early events in the evolution of Fungi, which warrants the relatively heavy sampling in these groups. In addition, microsporidia have been provisionally classified among the protists and were not previously thought to be members of the Fungi. Recent data suggest that they are nested among the basal lineages of Fungi (Keeling et al. 2000), but confirmation of their placement in the Fungi requires an accurate understanding of Chytridiomycota/Zygomycota relationships.

The sampling strategy of this project was designed in reference to recent higher-level phylogenetic studies in Fungi, which were reviewed by Taylor et al. (in press). These lists also include strains of all fungi that have been recently targeted for public genome projects (an initial list of fungal genomes proposed by the Whitehead Institute for the Fungal Genome Initiative is detailed at www-

genome.wi.mit.edu/seq/fungi). A summary phylogenetic tree showing the terminal clades from which individual species will be sampled and the approximate number of species to be sampled from each terminal clade is shown in Fig. 1. The detailed lists of species to be sampled in this project were developed in consultation with the Deep Hypha taxon discussion groups. Because of space limitations, the complete species lists are presented at the Deep Hypha web site. Criteria used to select species included 1) phylogenetic diversity; 2) availability of sequence data and nonmolecular characters; 3) culturability; 4) existence of high quality voucher material; 5) status as model systems, including subjects of genome projects; 6) economic and ecological importance; 7) nomenclatural type species (not all criteria apply to all species; order of criteria does not indicate relative importance); and, 8) availability of DNA, cultures, or fresh material from collaborators associated with this proposed project (see letters of support). The lists were developed as amendable documents that will be open to change as progress is made in sampling and new groups of Fungi are discovered.

Approximately 500 of the species necessary for this project are already on hand in the PI's laboratories. Of the remainder, about 400 are available in publicly accessible culture collections or herbaria, and 500 have been volunteered by collaborators (see letters of support and species lists at the Deep Hypha web site for details). Limited fieldwork will be necessary to collect fresh material of ~100 taxa, particularly microlichens, for which appropriate herbarium materials and cultures are unavailable. Outgroup choices for analyses within the Fungi will be determined in reference to the phylogenetic tree in Fig. 1. For analyses of the Fungi as a whole, multiple outgroups drawn from the Choanoflagellates, Animalia, and Mycetozoa will be selected, based on results of analyses of nuc-ssu rDNA and various protein-coding genes, which suggest that these groups are closely related to Fungi (Baldauf et al. 2000, Heckman et al. 2001, Keeling et al 2000, Wainwright et al. 1993). The AFTOL PIs will coordinate with zoologists and protistologists working in these groups to obtain sequences and/or DNA samples. An overview of sampling within the major groups of Fungi is given below.

C.4.2.2 Chytridiomycota/Zygomycota – Basal Fungal Lineages. The Chytridiomycota and Zygomycota comprise the most basal portion of the fungal tree of life. In this study, these Fungi are treated as a collective group of "Basal Fungal Lineages" containing at least ten (presumed) monophyletic groups (Fig 1). AFTOL sampling will include 100 species of Chytridiomycota and 200 species of Zygomycota from all 10 lineages. In addition, representatives of microsporidia will be included to further test their affinity to the basal lineages of Fungi. These taxa are being sequenced in the laboratories of P. Keeling and N. Fast, both of which are participants of Deep Hypha. Cultures for over 90% of these Fungi are already available to us (see supporting letters from K. O'Donnell, M. Powell, and J. Longcore). In addition, we will obtain unculturable and unusual species through experts who study such groups as necessary to round out our sampling. For example, specimens for phylogenetic studies are available through the International Collection of Arbuscular and Vesicular Arbuscular Mycorrhizal Fungi (INVAM, see letter from Joe Morton, Univ. W. Virginia). A summary of the Chytridiomycota-Zygomycota sampling is provided in Figure 1.

C.4.2.3 Basidiomycota. The 400 species of Basidiomycota to be sampled include 130 **Urediniomycetes**, 55 **Ustilaginomycetes**, and 215 **Hymenomycetes**. Ustilaginomycetes include roughly 6% of the described Basidiomycota, meaning that this group is somewhat overrepresented. Sampling in the Basidiomycota was designed based on phylogenetic trees and classifications by Swann et al. (1999, 2001) and Swann & Taylor (1995) (Urediniomycetes), Bauer et al. (2001) (Ustilaginomycetes), and Hibbett & Thorn (2001), Moncalvo et al. (2000, 2002), and Wells & Bandoni (2001) (Hymenomycetes). Sampling of basidiomycetous yeasts, which occur in all three subclasses of Basidiomycota, is based in part on the classifications of Fell et al. (2001) and Swann & Taylor (1995). Target species in the Urediniomycetes include all eight of the orders recognized by Swann et al. (2001), as well as members of the "*Naohidea* clade", "*Naiadella* clade", *Mixia osmundae*, and taxa that are of uncertain placement within the Agaricostilbomycetidae, Microbotryomycetidae, and Urediniomycetes. Target species in the Ustilaginomycetes include all nine of the orders recognized by Bauer et al. (2001) and several taxa that are of uncertain placement in the Ustilaginomycetidae. Target species in the Hymenomycetes include all seven orders of the heterobasidiomycetes sensu Wells & Bandoni (2001, plus Cystofilobasidiales), all of

the major (unranked) clades of homobasidiomycetes sensu Hibbett & Thorn (2001), and several additional groups of uncertain placement (e.g., the *Dendrocorticium* clade, *Gloeophyllum* clade, etc).
C.4.2.4 Ascomycota. The 800 species of Ascomycota to be sampled include 10 **Taphrinomycotina**, 25 **Saccharomycotina**, and 765 **Pezizomycotina**, a sampling mirrors the distribution of species diversity of the phylum. To avoid overlap with ongoing NSF-funded research (sect. C.5), the AFTOL sampling of Ascomycota will emphasize species, specimens, and/or genes not currently being sampled as part of other research programs. Taphrinomycotina includes four classes and five orders: Neolectomycetes, Pneumocystidomycetes, Schizosaccharomycetes, and Taphrinomycetes (Eriksson et al. 2001). AFTOL sampling of the Taphrinomycotina will focus on the Taphrinomycetes and target loci ATP6 and ITS for exemplars from the remaining classes that are included in other studies. While the Saccharomycotina contains one order, the Saccharomycetales, and ten families (Eriksson et al. 2001), recent studies are consistent with a large, undiscovered diversity of yeasts that will be detected through environmental sampling (M. Blackwell pers. comm., see letter of support). A well-resolved and robust phylogeny is needed so that current and future studies can place newly discovered taxa in the fungal tree of life and base their studies on a phylogenetic framework. A minimum of 2 exemplars will be sampled per family with a particular emphasis on the largest family Saccharomycetaceae. Pezizomycotina contains over 90% of the Ascomycota species diversity. Many species are only known to reproduce asexually and their integration into the classification of the Pezizomycotina, which is largely based on traits associated with sexual reproduction, is problematic; approximately 100 asexual species will be included in this sampling. Importantly, the Pezizomycotina includes more than 98% of Fungi that are lichenized. Lichenized Fungi account for approximately 42% of all described species of the Ascomycota and probably close to 50% of the known Pezizomycotina. Current classification of the Pezizomycotina includes 10 classes (Eriksson et al 2001), representing 10-15 separate lineages, with many of the interrelationships of the major groups (e.g., classes, subclasses, orders, etc.) still poorly resolved (Fig. 1; Spatafora 1995, Holst-Jensen et al. 1997, Berbee 1998, Liu et al. 1999, Gernandt et al. 2001, Lutzoni et al. 2001). The AFTOL sampling strategy will include multiple exemplars from all major classes, orders, and most families.

C.4.3 Molecular data acquisition. A total of seven loci will be sequenced, representing approximately 11,000 bp of nucleotide data per taxon. The loci that will be the focus of phylogenetic analyses include the nuclear (nuc) SSU rDNA (~1.7 kb), nuc-LSU rDNA (5' region ~1.4 kb), RPB1 (regions A-G ~2.5 kb), RPB2 (regions 3-11 ~2.5 kb), EF-1α (~1.4 kb), and the mitochondrial locus ATP6 (~0.8 kb). All loci were chosen based upon their known levels of appropriate nucleotide variation across the Fungi and the availability of preexisting data that have been used successfully in phylogenetic studies (e.g., Bruns et al. 1991; Berbee & Taylor 1993; Spatafora 1995; Berbee 1998; Liu et al. 1999; Baldauf et al. 2000; Moncalvo et al. 2000; James et al. 2000; Kretzer et al. 1999; Binder & Hibbett 2002; Gernandt et al 2001; Lutzoni et al 2001; Matheny et al 2002; O'Donnell et al 2001). ATP6 was also chosen to fulfill these specifications and to include a mitochondrial locus. It also embodies the desirable characteristic of lacking introns, unlike the majority of the fungal mitochondrial genes, which are replete with large and variable introns. ITS rDNA (~0.75 kb) will also be sequenced for the purpose of allowing fungal biologists the ability to approximate the placement of unknown specimens from environmental samples in the Fungal Tree of Life. ITS rDNA has been used successfully to identify unknown fungal species through relatively simple and highly efficient comparison (e.g., BLAST searches) to known fungal ITS sequences in Genbank (Bruns et al 1998, Vandenkoornhuyse et al 2002, O'Brian et al 2002).

PCR primers exist for all seven loci for some range of Fungi. (Specific information regarding fungal PCR and sequencing primers is available from the Deep Hypha website.) Universal and fungal specific rDNA PCR and sequencing primers are well developed and widely distributed among fungal systematists. PCR and sequencing primers have also been developed for RPB1, RPB2, EF-1a, and ATP6, but these genes are not nearly as well characterized as rDNA. Consequently, we anticipate that a significant effort will be required to design primers and refine PCR protocols. All of the PIs involved with the molecular aspects of AFTOL have considerable experience in developing molecular markers.

A large amount of sequence data that have potential utility for this project exists in publicly accessible databases (e.g., Genbank). However, much of these data cannot be used in this project because the data represent partial sequences for the regions we are targeting, and voucher specimens, cultures, and DNAs are not accessible. This situation demonstrates that much of the current data does not exist in a combinable and comparable form. We are proposing to collect and assemble approximately 11,000 bp of data for 1500 taxa, or 16,500,000 bp total. We will make every effort to include pre-existing data in this study. However, we estimate that only about 10% of the data required by the project currently exist in databases. Therefore, the total sequencing effort required for this study is about 14,850,000 bp.

The AFTOL project will focus on seven core loci because it is necessary to develop comprehensive, broadly combinable datasets for a large number of fungi. Nevertheless, we anticipate that this project will also involve the development of additional loci for fungal phylogenetics, both in the PI's laboratories and in collaboration with participating scientists. To facilitate this advancement, we will select a set of approximately 150 core taxa (10% of the total sampling) for which cultures will be established and a large amount of DNA will be stored and made available to fungal systematists who are interested in marker development. Dr. K. O'Donnell, curator of the NRRL culture collection, has agreed to provide long-term storage of these cultures. These well-characterized DNAs, distributed as "molecular exsiccatae", will constitute a significant resource for fungal phylogenetics; long-term -70°C storage of DNAs will be provided by R. Vilgalys. All new PCR and sequencing primers will be made publicly available.

All techniques associated with culturing Fungi, DNA extractions, PCR, and automated sequencing are in place and routinely performed in the laboratories of the PIs. Materials will be provided to the project from participating scientists as specimens, cultures, DNAs, PCR products, or some mix of these. Sequencing reactions will occur in the laboratories of the PIs in 96-well microtiter plates. These plates will be shipped to Duke University AFTOL sequencing facility where reactions will be run on an ABI 3700 automated sequencer. Raw sequences will be proofed for accuracy and subjected to a preliminary analysis to verify taxonomic identity (i.e., to confirm that the sequence is not that of a contaminant). After these quality-control measures, the sequences will be submitted simultaneously to GenBank, the AFTOL web-accessible database, and the collaborator who provided materials. For a more comprehensive treatment regarding collection of molecular data please refer to the Management Plan.

C.4.4 Non-molecular data acquisition. The volume of subcellular data is limited compared to the number of fungal species and the available molecular data. Generalizations based on this narrow database can be misleading. The available subcellular data are widely scattered in the literature and have not been compiled in a readily accessible format. To solve these problems, existing subcellular data will be compiled in a single database, and new data will be collected for selected taxa and characters.

Existing data on the following characters will be included in the database: spindle pole body (SPB) form and cycles, nuclear division (e.g., nuclear envelope changes, SPB-nuclear envelope interactions, chromatin and nucleolus behavior; Heath 1980, 1986; Swann et al. 2001), septum and septal pore organization (Kimbrough 1994; McLaughlin et al. 1995b; Wells & Bandoni 2001), meiospore and meiosporangium differentiation (Bellemere 1994; McLaughlin 1982; Clémençon 1997), selected cytoplasmic features (e.g., golgi apparatus, Spitzenkörper, microscala, colacosomes; Roberson & Fuller 1988; Swann et al. 2001), motile cell structure (Barr 2001), specialized cell types (e.g., paraphyses, cystidia; Clémençon 1997), haustorium-host interface (Bauer et al. 1997), and selected biochemical characters and wall chemistry (Bartnicki-Garcia 1970, 1987; Prillinger et al. 1993).

Acquisition of new subcellular data will focus on characters that are phylogenetically informative for broad groups of taxa, i.e., SPB, nuclear division, and septum. Assistance to other laboratories in obtaining structural data will not be limited to these characters, but may include other characters relevant to specific taxonomic groups. Acquisition of nuclear division and SPB data is inherently slow because of the number of cells that must be examined to observe all developmental stages. A maximum of about four of these studies per year may be possible. Other characters, such as those of the septum, can be acquired quickly. Target taxa for the nuclear division studies will focus on the largest gaps in the database: Glomales and

Trichomycetes where no data are available, and understudied clades in all phyla. In Chytridiomycota these data are available for 6 of 8 clades; in Zygomycota, for only 3 of 10 orders; in Ascomycota, for only 12 of 51 orders; and in Basidiomycota, for 19 of 35 orders and clades.

Methods for existing data: Working groups will be formed from the Non-molecular Participants Group to assess the data for each character, beginning with SPB, nuclear division, and septum. Additional characters will be prioritized for assessment. Character quality will be assessed to determine if it is adequate for evaluation or if new data are needed. Character illustrations will be chosen for the database, character stages defined and coded, and the data entered into the database.

Methods for new data: Cells will be grown, if culturable, or specimens collected. If cultured, preliminary tests will be made to determine the best way to find the desired stage (e.g., for nuclear division, hype are grown on agar-coated cover slips to allow cell selection of nuclei in partially fixed cells stained for fluorescence microscopy (Lü & McLaughlin 1995), or grown on cellophane over nutrient agar for freeze-substitution with cells selected later in thin plastic embedments (Swann et al. 1999). Other methods of cell preparation will include microwave-assisted specimen processing and high pressure freezing. Variations on these methods will be used for collected specimens. Cells will be serial sectioned, collected on slotted grids, and photographed with the transmission electron microscope. JPEG files of selected micrographs will be entered into the database.

C.4.5 Phylogenetic analyses. Because the data generated by this project will be available to the mycological community at large and because of the high number of labs involved (> 100 collaborators), the phylogenetic methods used will be very diverse, ranging from character based methods (maximum parsimony [MP] and likelihood [ML]) to distance based methods (e.g., neighbor-joining [NJ]). The size of the datasets will also vary greatly depending on the time at which the analyses will be conducted (e.g., year 1 compared to year 4 of this project), and on each researcher's interest (e.g., within family compare to the entire fungal kingdom). Four of the 5 PIs have conducted extensive phylogenetic analyses on datasets with high number of taxa for each of the four respective phyla within the kingdom Fungi (e.g., Hibbett et al. 2000, James et al. 2000, Kauff and Lutzoni 2002, Lutzoni et al. 2001, Moncalvo et al. 2000 and 2002, O'Donnell et al. 2001, Platt and Spatafora 2000) and are all aware of the computational and methodological limitations imposed by such data matrices. The largest data matrix that will be analyzed as part of this proposed project will be at least 1500 species by six loci (about 10,000 bp, excluding ITS).

The first analytical challenge (other than alignment, see Management Plan) will be to choose a test for detecting conflicts among data partitions that if combined would lower phylogenetic accuracy (Barker and Lutzoni 2002, Bull et al. 1993, Cunningham 1997a, 1997b). Even if many phylogeneticists agree with the principles described in Bull et al. (1993), that data partition homogeneity should be tested before combining them, the question remains which test should be chosen (Bullet al. 1993, Farris et al. 1995a, 1995b, Huelsenbeck and Bull 1996, Huelsenbeck et al. 1996, Rodrigo et al. 1993, Mason-Gamer and Kellogg 1996, Waddell et al. 2000), when all the most readily available methods are known to be poor predictors as to whether or not combining two specific data partitions is likely to lower phylogenetic accuracy (Barker and Lutzoni 2002, Dolphin et al. 2000, Dowton and Austin 2002, Cunningham 1997a, 1997b, Graham et al. 1998, Lutzoni 1997, Lutzoni and Barker 1999, Lutzoni and Vilgalys 1995, Mason-Gamer and Kellogg 1996). For such large data sets, we plan to use neighbor-joining with non-parametric bootstrap and use 70% as the threshold value for detecting incongruence that is likely to decrease phylogenetic accuracy if data partitions are combined. Using this method, if a set of taxa part of a monophyletic group for one data partition has a bootstrap value of ≥ 70% and the same set of taxa is non-monophyletic with support value ≥ 70% for another data partition, the problematic sequences will be examined to determine if this is due to an error during the molecular work, or due to a flaw in the analysis such as an inappropriate distance correction. If the former is the cause, the sample will be resequenced for the problematic locus. If a sensitivity analysis using different measures of distance shows that the analysis was the cause for this conflict the analysis will be corrected. If the conflict is real (e.g., due to lineage sorting or hybridization) the problematic taxa will be removed and replaced by a different taxa.

As part of a CAREER grant to F. Lutzoni (DEB-0133891) simulations will be conducted in collaboration with Stefan Zoller to develop new methods to address this issue.

The second challenge will be to conduct phylogenetic searches and estimate confidence levels for the resulting phylogenies derived from the combined dataset. For character based methods such as ML and MP the Ratchet search algorithm will be used (Nixon 1999) using PAUPRat with PAUP* (Swofford 1998, Sikes and Lewis, pers. comm.). We will also explore the new program TNT, designed specifically by Goloboff, Farris and Nixon for analyzing large data sets. Neighbor-Joining, implemented in PAUP*, will also be used. Confidence for specific tree partitions will be estimated using parametric bootstrapping. If MrBayes (Huelsenbeck) or other programs implementing Bayesian Markov chain Monte Carlo (BMCMC) tree sampling is able to handle such large dataset (and can be run in parallel), BMCMC will also be used to estimate confidence levels associated with each node. All these large phylogenetic analyses will be conducted in collaboration with Jack da Silva at the North Carolina Supercomputing Center where we have access among other supercomputers to an IBM RS/6000 SP with 720 processors (see letter of support by Jack da Silva). Even if programs such as PAUP* are not programmed to run in parallel yet, it is relatively simple to partition the job into multiple smaller jobs and distribute these jobs to these 720 processors. Finally, specific analyses will be conducted to integrate non-molecular characters with molecular characters (Lutzoni and Vilgalys 1995).

C.4.6 Development of web-accessible resources for fungal phylogenetics. All five PIs have experience with database management and web-based dissemination of results (four of the five PIs are also co-authors for the *Tree of Life* Web Project, http://tolweb.org/tree/phylogeny.html). A vehicle for web-based coordination and dissemination of results from this study already exists, through the Deep Hypha RCN website that is managed by JWS. This web site will be expanded to provide access to results from this study. We will enhance the utility of this site through hyperlinks to other on-line databases for mycology and molecular systematics (GenBank, TreeBASE, FungalWeb, Systema Ascomycetum, Morphobank). Below is a brief synopsis of web-based resources that will be developed in the AFTOL project:

1. **Information about primers and genes**. We have already established links for pages discussing the utility of different genes and primer sets for phylogenetic analysis, along with lists of taxa and specimens which we propose to sequence for this study.
2. **Data uploading**. The database management plan for this project specifies release of sequence data onto the web server as soon as each sequence is assembled and proofed. Participating outside labs will also be allowed to contribute sequences to the database on-line. Each PI's lab will be responsible for proofing sequences prior to release. The sequence database server will also include additional sequence data for taxa that are not part of this project (updated weekly using a "web-bot" that automatically scans the NCBI database). As new genes are developed, databases for each will also be added to the site.
3. **Downloading utilities.** A download area on the web site will allow users to download individual sequences and alignments, with links to other databases, such as TreeBase (http://www.treebase.org).
4. **WWW-based tools for sequence identification.** One goal of this study is the development of tools for the identification of unknown fungal sequences on the Internet, especially ITS sequences that are proving to be very useful for species-level identification. Users of the AFTOL web site will be able to: i) separate sequences into their conserved and variable components; ii) perform BLAST searches; iii) automatic alignment and phylogenetic analysis. Such a system would be of great value to a wide variety of end-users including systematists, ecologists, plant pathologists, and microbiologists. We already have support for developing this part of our project through the resources of Duke Center for Bioinformatics and Computation Biology (see supporting letters from Kim Johnson and John Harer).
5. **Phylogenetic trees.** Phylogenies will be posted on the site with other new results during the course of the investigation. Interactive hyperlinks will be developed for exploration of fungal evolution within a phylogenetic context. For example, phylogenetic reconstruction of non-molecular characters may be mapped onto each tree, with further links to more detailed discussion of each character, taxon, etc. Links to other phylogenetic studies/sites (e.g., Tree of Life web page) will be included here as well.

C.4.7. Data release, publication, and dissemination of results. As noted above, all data will be released via the world-wide web as soon as they have been generated. The results of analyses will be disseminated by traditional means (journal articles, presentations, etc) as well as the web-based resources described in the previous section. In addition, each of the PIs will contribute to updating the Tree of Life Web Project pages for Fungi (http://tolweb.org/tree/phylogeny.html).

Policy regarding authorship: All persons who contribute material to the AFTOL project will be invited to co-author papers that present new data derived from the material that they donated. In addition, one or more manuscripts that synthesize the major finding of the AFTOL project will be published. It is anticipated that such a summary publication(s) will be produced toward the end of the funding period and that authorship of this publication will consist of all participating scientists in Deep Hypha and AFTOL.

C.4.8. Outreach and K-12 education. The AFTOL project will promote the inclusion of fungal biology in K-12 education through two mechanisms:

1. *Summer workshops in fungal biology for high school teachers*. Each of the four institutions involved in this project will host a two-day workshop every summer for local high school teachers, which will provide information about the basic biology and ecological roles of Fungi, economic and human health impacts of Fungi, and recent advances in fungal phylogenetics. The latter will include an overview of methods in phylogenetic analysis, as well as current hypotheses regarding higher-level relationships of Fungi (relative to plants, animals, and other eukaryotes). To the extent possible, the PIs will arrange to have these workshops count toward professional development credits, which are required for annual teacher recertification in the public school systems in all of the states of the PI's institutions (a model for such activity is being implemented in DEB 0128925, to DSH).

2. *Development of web-accessible resources for fungal biology for K-12 educators*. A "Fungal K-12" website will be created with information and resources that will enable K-12 educators to develop lessons involving Fungi. The topics discussed above will be presented, as well as suggestions for simple laboratory exercises involving Fungi, and suggested literature. Links to other fungal biology web sites of interest to K-12 educators and students will be included. The Fungal K-12 website will be developed at Clark University (see DSH Budget Justification) and linked to the Deep Hypha website.

C.5 RELATIONSHIP TO OTHER NSF-SUPPORTED ACTIVITIES IN FUNGAL SYSTEMATICS.

C.5.1 Relationship to Deep Hypha. Deep Hypha supports two annual meetings of fungal systematists. The AFTOL project depends on a high degree of community participation, and it will be greatly facilitated by Deep Hypha. Indeed, all participants in Deep Hypha have already been consulted regarding the development of the present proposal, and many have volunteered to participate by supplying materials, advising on taxon selection, etc (see letters of support). The AFTOL project will continue to use the Deep Hypha meetings, web site, and e-mail list to consult with fungal systematists about research priorities, as well as to present recent research developments. In addition, the AFTOL PIs will offer brief workshops as part of Deep Hypha, which will focus on methods of phylogenetic analysis, new loci for fungal systematics, and other subjects that are designed especially for students and beginning researchers in fungal phylogenetics. Thus, Deep Hypha and AFTOL will be highly integrated. However, because Deep Hypha has already been funded by the NSF, no funds are requested in this proposal to support it. The final Deep Hypha meeting is scheduled for summer, 2005, which means that there will be no Deep Hypha meeting in the last year of the AFTOL project. JS is a Co-PI of Deep Hypha, DSH is a member of the Deep Hypha Steering Committee, and all PIs of the present proposal are participants in Deep Hypha.

C.5.2 Relationship to NSF-funded work in progress elsewhere. There are sixteen "currently active" projects in fungal systematics that are funded by the Systematic Biology Program (as of 5/5/02), excluding ongoing projects in the PI's laboratories (see below) and Dissertation Research grants. Projects that are most closely related to the proposed AFTOL project include studies on Chytridiomycetes (DEB 9978094, PI Powell), Trichomycetes (DEB 0108110, PI Lichtwardt), Sordariales (DEB 0118695, PI Huhndorf), Sarcoscyphineae (DEB 9521944, PI Pfister), Cladoniaceae (DEB 9712484, PI DePriest), Agaricales (DEB 0118776, PI Desjardin), Agaricales and Aphyllophorales (DEB 9978011, PI Petersen),

and Russulaceae (DEB 9974018, PI Miller). All of these projects focus on individual clades, with intensive sampling. The AFTOL project will facilitate these other investigations by clarifying sister-group relationships and developing new loci for phylogenetics. Many of the PIs of the currently active fungal systematics projects have indicated that they will participate in AFTOL (see letters of support).

C.5.3 Relationship to ongoing work in the PI's laboratories. DSH currently has one NSF grant to study higher-level phylogeny of homobasidiomycetes (DEB 9903835). This project, which will end in August, 2002, has already generated most of the nuclear rDNA sequences necessary for the homobasidiomycetes, as well as about 25 RPB2 sequences. However, no RPB1 or EF-1α sequences, or any sequences from Urediniomycetes and Ustilaginomycetes, have been generated. A new project on relationships of cyphelloid and aquatic homobasidiomycetes will begin in DSH's lab in September, 2002 (DEB0128925), but this project will have minimal overlap with the AFTOL project. RV currently has two ongoing NSF grants which provide useful models for the high-volume sequencing strategies that we will employ in this study: 1) "Phylogenetic Systematics of the Agaricales" (RV with JM Moncalvo, DEB-0076023) is funded through June 2003, and involves significant collaboration with the larger taxonomic community to establish a phylogenetic database, including study of over 1000 diverse spp. of mushrooms, puffballs, and other euagarics; 2) The Duke Forest Mycological Observatory (DFMO) is a new project (MCB-0084207) that applies molecular systematics for the study of fungi through analyses of environmental DNA sequences. In addition to laboratory methods, the DFMO also provides information about the extent of "undiscovered" diversity in fungi (and other microorganisms) that is highly relevant to this proposal. DJM is currently involved in analysis of Urediniomycetes, both structural and molecular, in cooperation with E. Frieders and L. Szabo, and planning is underway to expand the analysis of the Uredinales. He is also participating in studies of ectotrophic mycorrhizal (ECM) fungi in oak savannah at a nearby LTER site which involves phylogenetic analysis of ECM community structure. The high species diversity at this site will supply us with a readily accessible source of field material for structural analysis. JWS currently has one NSF grant to study familial-level phylogeny of the Clavicipitaceae (Ascomycota, Hypocreales) (DEB-0129212). This grant has minimal overlap with the proposed research. It will generate SSU rDNA, LSU rDNA, and EF-1α the sequence data for 14 taxa to be included in the Ascomycota sampling, but it will not generate any of the RPB1, RPB2, ATP6, or ITS sequences. FML currently has a NSF-CAREER grant, the goal of which is to develop a multigene phylogeny for the Ascomycota (DEB-0133891). That project will sample approximately 150 taxa for five of the same genes to be sequenced as part of this research. These 150 taxa are NOT part of the 800 taxa that will be sampled as part of this project. Rather, the sampling strategy in DEB-0133891 and the one proposed here were developed to complement one another and to provide a more comprehensive sampling of the major groups that house more than 35,000 species of Ascomycota, especially inoperculate discomycetes of the Leotiomycetes, orders and families *incertae sedis*, and asexual species of unknown phylogenetic affinity.

C.6 SIGNIFICANCE

Fungi represent one of the most important clades of eukaryotes. They perform a large percentage of the nutrient cycling in ecosystems, cause diseases of plants and animals (including humans), and produce novel secondary metabolites that are used in the development of pharmaceuticals. A comprehensive fungal phylogeny will benefit diverse disciplines, from community ecology and plant pathology, to medical mycology and drug discovery. A barrier to the development of a comprehensive fungal phylogeny has been the lack of combinable and comparable multi-gene sequence data. AFTOL will dramatically increase the amount of multi-gene sequence data for fungal systematics and will transform the field from one based primarily on rDNA to one that embodies a more extensive survey of the genome. In addition, the data generated by AFTOL, especially ITS data, will promote the discovery and detection of Fungi using molecular techniques, such as microarray methods (see letters from J. Taylor and D. Geiser). Finally, through training of postdocs and graduate students, and outreach activities directed at K-12 educators, AFTOL will contribute to training the next generation of fungal systematists, promote inclusion of fungal biology in K-12 education, and facilitate collaboration among mycologists.

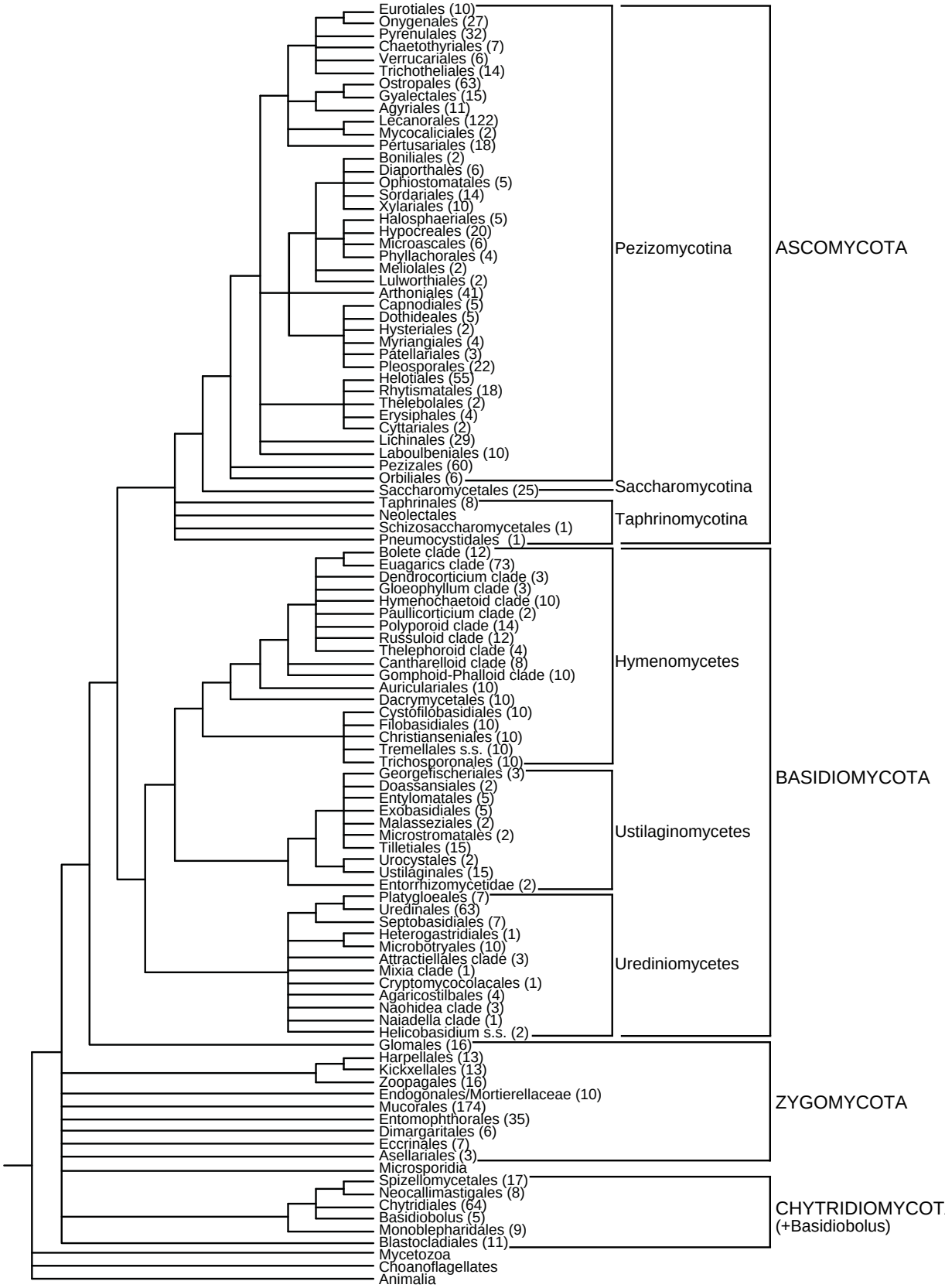

Fig. 1. Higher-level phylogenetic relationships of Fungi and overview of taxon sampling. This tree represents current hypotheses of fungal phylogenetic relationships, based on published molecular phylogenetic analyses (for references, see sects. C.3, C.4.2). Numbers in parentheses following names of terminal taxa indicate approximate numbers of species to be sampled in each group. Detailed lists of taxa, including taxa to be sampled for non-molecular characters, are posted at the Deep Hypha web site. About 100 mitosporic species and several meiosporic taxa of uncertain placement, which will be sampled, are not represented in this tree. Several taxa that are being studied by other projects are present on the tree, but no data will be generated for these taxa in this study (Neolectales, Microsporidia, and outgroup taxa).