ARTICLE

Research and teaching with the AFTOL SBD: an informatics resource for fungal subcellular and biochemical data

T. K. Arun Kumar[1], Meredith Blackwell[2], Peter M. Letcher[3], Robert W. Roberson[4], and David J. McLaughlin[5]

[1]Department of Botany, The Zamorin's Guruvayurappan College, Calicut, Kerala 673014, India; corresponding author e-mail: tkakumar@gmail.com

[2]Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA [3]Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA

[4]School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

[5]Department of Plant Biology, University of Minnesota, St Paul, MN 55108 USA

Abstract: The Structural and Biochemical Database (SBD), developed as part of the US NSF-funded Assembling the Fungal Tree of Life (AFTOL), is a multi-investigator project. It is a major resource to present and manage morphological and biochemical information on *Fungi* and serves as a phyloinformatics tool for the scientific community. It also is an important resource for teaching mycology. The database, available at http://aftol.umn.edu, includes new and previously published subcellular data on *Fungi*, supplemented with images and literature links. Datasets automatically combined in NEXUS format from the site permit independent and combined (with molecular data) phylogenetic analyses. Character lists, a major feature of the site, serve as primary reference documents of subcellular and biochemical characters that distinguish taxa across the major fungal lineages. The character lists illustrated with images and drawings are informative for evolutionary and developmental biologists as well as educators, students and the public. Fungal Subcellular Ontology (FSO), developed as part of this effort is a primary initiative to provide a controlled vocabulary describing subcellular structures unique to *Fungi*. FSO establishes a full complement of terms that provide an operating ontological framework for the database. Examples are provided for using the database for teaching.

INTRODUCTION

Before there were DNA markers, reconstructions of fungal evolutionary history relied heavily on structural and biochemical data. Today, rapidly accumulating molecular databases reveal new or uphold old taxonomic hypotheses but upset some hypotheses that were based solely on morphology. Problems arise with molecular and biochemical characters primarily because it is not always possible to distinguish homoplasy and because the characters cannot be polarized. Structural characters, however, remain important to support certain lineages in phylogenetic analyses and to provide data when DNA is not available. Phylogenies can be used to reevaluate characters in order to recognize homologies (Celio *et al.* 2006). Recent recognition of the importance of combined morphological and molecular data for improved phylogenetic analyses highlights the continued significance of structural information (Kumar *et al.* 2012, Padamsee *et al.* 2012, Healy *et al.* 2013). In addition, biochemical pathways such as membrane sterols may be useful in distinguishing certain fungi such as powdery mildews from other ascomycetes or rust fungi from agarics (Weete *et al.* 2010).

The Structural and Biochemical Database (SBD, https://aftol.umn.edu), developed as part of the US National Science Foundation-funded Assembling the Fungal Tree of Life (AFTOL) project, is an online resource for researchers to compile fungal subcellular and biochemical characters and character state data for phylogenetic analyses, including tree and ancestral state reconstructions. Subcellular and biochemical data were collected from published studies and new research, their quality and taxonomic informativeness were assessed as described in Celio *et al. (*2006), recorded in Excel files and finally entered into the database. Quality images to support the data were uploaded. Characters and character states (e.g. Nuclear Division, Septum/Pore Cap, Motile Cells, Sterol) have been assembled, character coded, illustrated and compiled as lists in downloadable PDF format. Data accompanied by images and diagrams from published and unpublished sources are useful for understanding the complexity of the characters. The illustrations used are either from published literature or generated from drawings using Adobe Photoshop and Adobe Illustrator programs. With the help of the SBD, data matrices can be generated to test or compare analyses done independently or done in combination with molecular data.

Fig. 1. Title page of the AFTOL Structural and Biochemical Database with links to the information available in the database.

The SBD also contains the Fungal Subcellular Ontology (FSO), a controlled vocabulary that describes the detailed subcellular structures unique to fungi. A full complement of terms was established to provide an ontological framework for the database (Kumar *et al.* 2011). Ontologies are useful for communication across databases, such as from ours to databases being developed by other scientists. The ontology (FSO) developed for the SBD has the terms organized under the commonly used 'is_a', 'part_of' class relationships in OBO-Edit format, edited using OBO-Edit 1.101. FSO in OBO-Edit format is hosted at: https://aftol.umn.edu/ontology.

Highlights of the SBD
How can you use the SBD? : The SBD is a phyloinformatics tool that includes published as well as new data on

subcellular and biochemical characters (https://aftol.umn.edu) supplemented with images and literature links (Fig. 1). *You may view the species in the database and determine their taxonomic status* (Fig. 2). Ultrastructural data are compiled under the headings of Motile cells, Nuclear division, and Septum/Pore cap. Biochemical data on sterols also have been incorporated. Compilation of data on hyphal tip organization, meiosporangia morphology, cystidial morphology and cell wall carbohydrates is in progress but not yet available.

Available structural and biochemical data are displayed on a page for each species. The cell type (e.g. diploid mycelium) of the organism for which the data are available is specified. This information is followed by the characters and coded character states (0, 1, 2, …). The sources section provides the data source (unpublished AFTOL data are indicated as

Fig. 2. An example from the database showing certain ascomycete taxa ("Taxonomy") for which data can be retrieved.

"AFTOL data") and voucher details. Images are included for characters whenever possible when we were provided blanket copyright permissions from cooperative publishers. Restricted use by certain journals has been a problem in compiling the database, and we strongly encourage authors who wish their data to be made available to the scientific community to choose journals with appropriate scientific policies on reuse of images or to retain copyright to their images. We applaud the open access concept.

The database provides automatic compilation of fungal subcellular and biochemical character state data in NEXUS format for independent and combined (e.g. with molecular data) phylogenetic analyses for completed data sets. The **N**EXUS files can be generated based on searches using: (1) Taxonomy, (2) Character, or (3) Specific character state.

The "Character lists" are primary reference documents for important subcellular and biochemical characters that help to distinguish taxa across the major fungal lineages. Illustrated character lists with images and drawings are an invaluable resource for educators and students as well as evolutionary biologists (Fig. 3). These lists can be downloaded from the database as PDFs (https://aftol.umn.edu/glossary).

Although we expect the SBD to aid in genomic analyses, integrating structural and biochemical characters with genomic data presents challenges. The incompleteness of structural studies in fungi, which is revealed by the precise comparisons required in determining character states for entry into the SBD, makes character evolution studies difficult, for example, in comparing septal pore characters between *Ascomycota* and *Basidiomycota* with limited studies in the *Taphrinomycotina*. Marker genes from genomic analyses, however, may aid in character definition as has been shown with *hex-1* genes for Woronin bodies that separate them from similar appearing microbodies or other structures associated with ascomycete and basidiomycete septal pores (Dhavale & Jedd 2007, Healy *et al.* 2012). Also, because of disparate rates of evolution and convergence among molecular and structural characters, integration of molecular and structural data may be limited or restricted to individual lineages.

Biochemical data often are difficult to interpret when we rely on gene expression to determine presence or absence of a character. For example, sterol data depend on determining presence or absence of a sterol that may be compounded by synthesis of multiple sterols or multiple uses of an enzyme in different pathways. Lysine synthesis pathways have been used to separate fungi from other organisms, including oomycetes. *Fungi* are said to have the *a*-aminoadipate (AAA) lysine synthesis pathway while most other organisms have a diaminopimelate (DAP) pathway. The genes for two key enzymes (AAR and lysA), respectively markers of the AAA and DAP pathways, however, have spotty distributions among eukaryotes and, moreover, both genes occur in the fungi that have been assayed (Torruella *et al.* 2009). Eventually complete genomes and analysis of entire pathways may help to lessen the perceived complexity of such distributions. In another example "shared signatures" from genomic studies have helped to unite *Microsporidia* and *Rozella* within the fungi (James *et al*. 2013). However, it should be noted that flagellar proteins present in zoosporic fungi and *Naegleria* are lost in ascomycetes and basidiomycetes; they also are

absent in microsporidians, indicating convergence is possible even when entire genomes are examined (James *et al.* 2013).

The Fungal Subcellular Ontology (FSO), also developed as part of the AFTOL2 project (https://aftol.umn.edu/ontology), uses controlled vocabularies to describe subcellular structures unique to fungi (Kumar *et al.* 2011). The ontology provides uniform definitions for characters included in the database and describes relations between them in a formal way. An ontology-based design enhances reuse of SBD data from a diverse group of independent electronic biological and genetic databases.

THE FUTURE

We welcome additions from other researchers and hope that the database will continue to grow. A contributor can directly access the template Excel files from the database and add data. Once the data are put into the Excel files, the contributor can contact the SBD administrator for upload of the data. The administrator will review the data file and send it to expert AFTOL curators for in-depth review. Accepted data files will be given unique SBD accession numbers. Quality images with copyright permissions greatly enhance and support the data entered, and these can be uploaded to the database by the administrator. Unpublished images that supplement published data also are desirable. Images should be scanned at a high resolution (700–1200 dpi) and saved in TIFF or JPEG formats. JPEG images (without compromised quality) are the preferred format. Source acknowledgements may be given in a separate box provided for adding legends or optionally on the image itself (in Arial 10 pt. font). For published or unpublished images an indication of the copyright holder and license are needed. The preferred licenses are in the public domain, Creative Commons Attribution-Share Alike (see Flickr), or another open content license. If this sounds overwhelming, don't be concerned because the administrator will walk you through it.

THE SBD FOR TEACHING EXERCISES

The SBD can be used by students at many levels for learning more about fungal morphology and biochemistry. Students with phylogenetic expertise can make trees using the data to assemble data matrices based on subcellular characters available in the SBD. Ancestral character state reconstructions and mapping of evolutionary change of characters can be useful exercises for students. If the students have experience using phylogenetics software, they can download Mesquite software (http://mesquiteproject.org/mesquite/mesquite.html) to analyze comparative data about fungi on a phylogenetic tree. Alternatively, students with little phylogenetic experience can be provided by the instructor with a tree (Fig. 4) on which characters can be mapped by hand to learn character variation within the fungi.

For example, for tracing character evolution, lay out the exercise in a numbered series for each step using a specific example, as in the *Orbilia* analysis (Kumar *et al.* 2012). Include the procedure for downloading Mesquite. Character
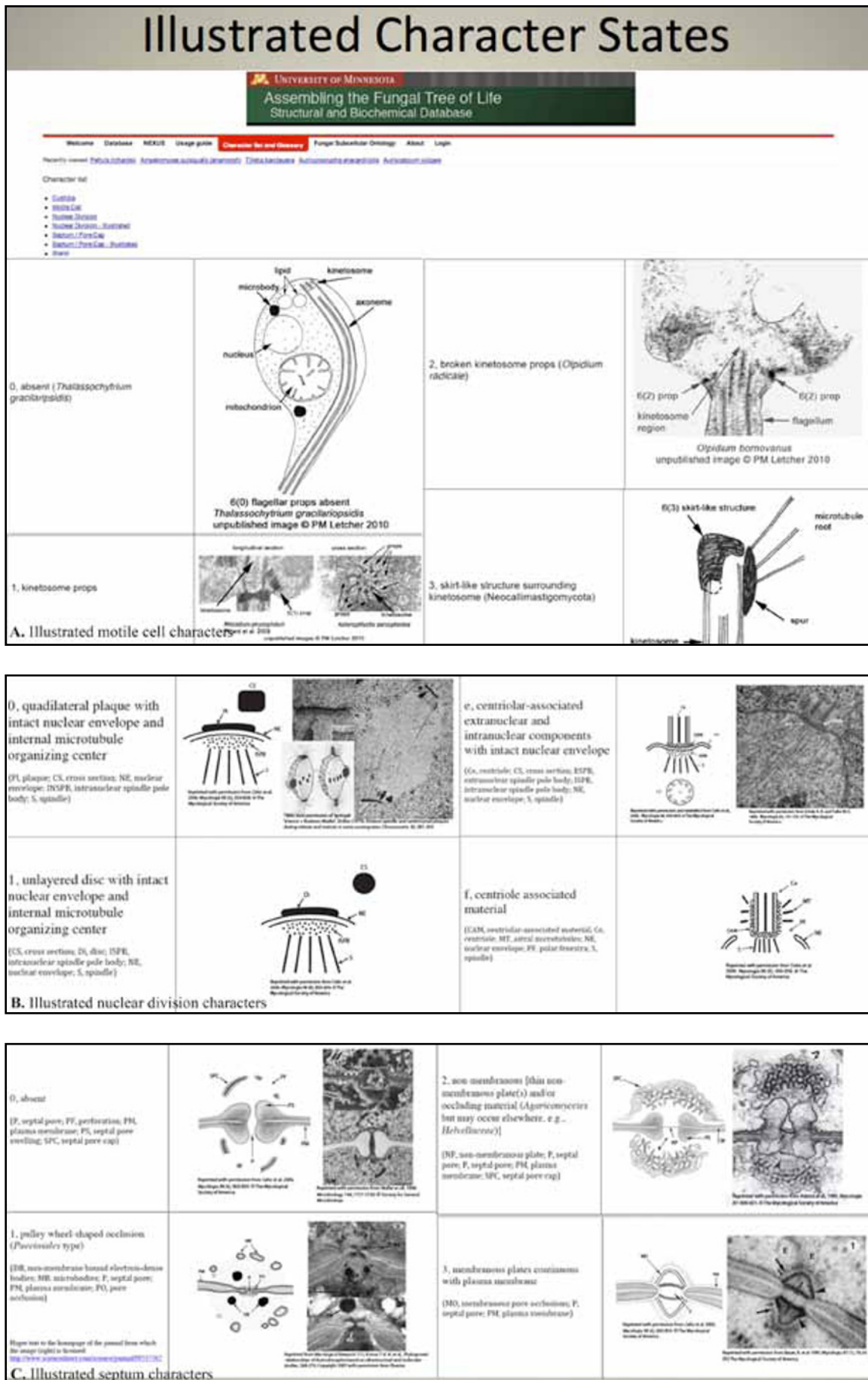
Fig. 3. Examples of the type of detailed data available. **A.** The complex flagella apparatus of motile cells, **B.** Nuclear division including spindle pole body. **C.** Septal pore characters and associated structures.
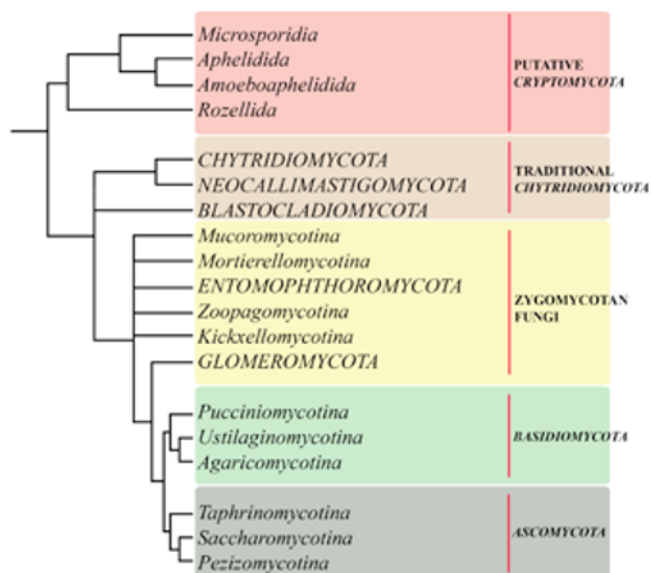
Fig. 4. Cladogram showing phylogenetic relationships of the major lineages of Fungi.

sets that could be used in tracing character evolution in other fungi are motile cells, septa, nuclear division, sterols or apical organization.

Students at many levels may enjoy a treasure hunt. Use the character state lists to begin to find the answers to the following:

- What is a spindle pole body (SPB)? Where did you find the answer?
- Motile fungi have centrioles at the spindle poles during nuclear division. What fungi suggest that the centriole was lost during the evolution of the SPB in organisms that lack motile cells?
- What changes occur in SPB form in *Ascomycota* or *Basidiomycota*? Do any of the SPB types suggest a link between these phyla?
- Which groups of fungi have flagella?
- Which group has more than one flagellum?
- Do multiperforate septa indicate phylogenetic relatedness? Why or why not?
- What is unusual about the septal pore organization of ascogenous hyphae or asci in some *Pezizomycotina*?
- What septal pore character supports the idea of a common origin (sister relationship) of *Ascomycota* and *Basidiomycota*?
- What changes occurred in the margins of the septal pores in *Basidiomycota*, or in zygomycetous fungi?
- Where are Woronin bodies found? What morphological variation do they show?
- In which fungi are septal pore caps found? How many kinds can be recognized?
- Presence of ergosterol cell membranes is often listed as a trait of all fungi. Is this an accurate assessment?
- How would you use sterols to distinguish the powdery mildews (*Erysiphales*) from other fungi that may occur on the surface of leaves?

- Why are studies that rely on ergosterol assays to detect *Glomeromycetes* in soil not accurate?
- Can you think of more characters that could be added to the database? Your textbook or reliable internet sites could help with ideas. Remember that you may distinguish taxa at any level.

A primary aim of the project has been to collect hard to obtain data in a useful format. We hope that you will use and provide feedback on the SBD and the ontology.

ACKNOWLEDGEMENTS

REFERENCES

Celio GJ, Padamsee M, Dentinger BTM, Bauer R, McLaughlin DJ (2006) Assembling the fungal tree of life: constructing the structural and biochemical database. *Mycologia* **98**: 850–859.

Dhavale T, Jedd G (2007) The fungal Woronin body. In: *The Mycota.* Vol. 8. *Biology of the Fungal Cell* (Howard RJ, Gow NAR, eds): 87–96. Berlin: Springer-Verlag.

Healy RA, Kumar TKA, Hewitt DA, McLaughlin DJ (2013) Functional and phylogenetic implications of septal pore ultrastructure in the ascoma of *Neolecta vitellina*. *Mycologia* **105**: 802–813.

James TY, Pelin A, Bonen L, Ahrendt S, Sain D, Corradi N, Stajich JE (2013) Shared signatures of parasitism and phylogenomics unite *Cryptomycota* and *Microsporidia*. *Current Biology* **23**: 1548–1553.

Kumar TKA, Crow JA, Wennblom TJ, Abril M, Letcher PM, Blackwell M, Roberson RW, McLaughlin DJ (2011) An ontology of fungal subcellular traits. *American Journal of Botany* **98**: 426–438.

Kumar TKA, Healy R, Spatafora JW, Backwell M, McLaughlin DJ (2012) *Orbilia* ultrastructure, character evolution and phylogeny of *Pezizomycotina*. *Mycologia* **104**: 462–476.

Padamsee M, Kumar TKA, Riley R, Binder M, Boyd A, Calvo A, Furukawa K, Hesse C, Hohmann S, James TY, LaButti K, Lapidus A, Lindquist E, Lucas S, Miller K, Shantappa S, Grigoriev I, Hibbett DS, McLaughlin DJ, Spatafora JW, Aime MC (2012) The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. *Fungal Genetics and Biology* **49**: 217–226.

Torruella G, Suga H, Riutort M, Peretó J, Ruiz-Trillo I (2009) The evolutionary history of lysine biosynthesis pathways within eukaryotes. *Journal of Molecular Evolution* **69**: 240–248.

Weete JD, Abril M, Blackwell M (2010) Phylogenetic distribution of fungal sterols. *PLoS ONE* **5**(5): e10899. DOI:10.1371/journal.pone.0010899.