

# Mycena News



The Mycological Society of San Francisco October 2008, vol. 59:07

## Speaker for the October 21 MSSF Meeting



Dennis Desjardin, PhD

### Mushrooms of the Islands at the Center of the World

Dr. Desjardin will present wonderful data on his latest expedition to Sao Tome and Principe, the country in west Africa south of Cameroon and west of Gabon—as close to latitude 0° and longitude 0° as you can get without getting too wet!

Continued on page 7

## Inside This Issue...

MycDigest.....	1
Speaker for October 21, 2007.....	1
President's Post.....	2
Announcements.....	2-3
MSSF Scholarship Info.....	3
About Drying and My Dryer.....	4
NAMA Foray Report.....	6
Fall Recipe: Porcini Pasta Sauce.....	7
MSSF October Calendar.....	10

## MycoDigest: Fungal Genomes

Else C. Vellinga

The publication of the human genome in 2001 was a milestone in our understanding of the human genetic make-up. The data shed light on the number of protein-coding genes (20–25 thousand, far fewer than expected), their position and composition, and the rest of the genome, which lies between the genes, and whose significance is still obscure. The data are a treasure trove for biomedical science. For example, they have been widely used to hunt for the genes that cause, or make us susceptible to, particular diseases.

What exactly is a genome sequence? It is the order, like letters in a text, of four different bases (bases being a particular kind of molecule) in a chain of millions, an order which scarcely varies from one individual to another in the same species. We say that this order is the genetic code of the species and different orders make different species. But the order is more than a name: it is a set of specifications for making the myriad chemical building blocks of life. The bases form the rungs of a twisted ladder which is the structure of the DNA molecule (the “sides” of the ladder, famously called the double helix, are made up of sugars and phosphates).

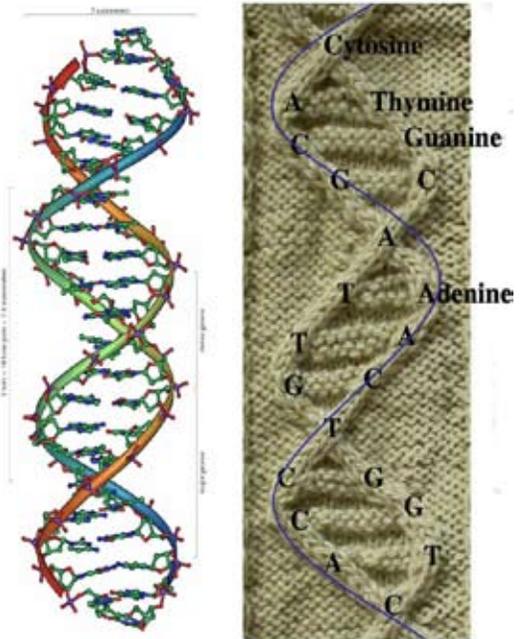


Fig. 1. Left: the DNA molecule, with helices connected by base pairs or “ladder rungs.” Right: a knit rendition of the DNA molecule’s double helix; the sequence along the line reads as: CACGT (behind the A) and AACTCCAG.

Left: image courtesy of Michael Ströck, [http://commons.wikimedia.org/wiki/Image:DNA\\_Overview.png](http://commons.wikimedia.org/wiki/Image:DNA_Overview.png). Right: DNA cable pattern by June Ishira, knit by Else Vellinga.

Continued on page 8

MycoDigest is a section of *Mycena News* dedicated to the scientific review of mycological information.

## MycoDigest continued

Does this mean that there are two orders, corresponding to the two sides of the ladder? Not really. Eukaryotes have four different DNA bases (adenine, cytosine, guanine, and thymine), which are paired (A-T and C-G) with one member of each pair on a different side (see fig.1). Because of this correspondence, the order of bases on one side of the ladder can be read from the bases on the other side. A part of the ladder that codes for the making of a particular protein or enzyme is called a gene; three bases in a row code for one amino acid, and many amino acids (often hundreds) make up proteins and enzymes. There are patterns in the code that mark the beginning and end of genes and the intervening regions are called “non-coding,” though their significance is not understood. DNA is organized in chromosomes in the cell nucleus, but also in organelles within the cell, like mitochondria, which were once bacteria.

Humans were by no means the first species to have their genome sequenced. Bacteria, with their small genomes, were the forerunners, and the first eukaryotic (non-bacterial) organism was a fungus: baker's yeast (*Saccharomyces cerevisiae*) (fig.2). Next came three other genetic models: the first multicellular organism, the nematode *Caenorhabditis elegans* (colloquially called *C. elegans*) (1998); the fruit fly *Drosophila melanogaster* (2000); and the plant model, *Arabidopsis thaliana* (also in 2000). The *Homo sapiens* genome was ready in 2001.

The publication of the yeast genome in 1996, with the title *Life with 6000 Genes*, is still a very interesting read. The functions of many of its genes were not known at that time. In retrospect, the yeast genome seems very compact and low in non-coding regions. The work on that first sequencing project took many years, involved 600 scientists (only 16 of whom became co-authors of the paper), and many institutions worldwide. At that time it was daringly estimated that the human genome sequence would be ready in 2005. The invention of different, faster sequencing techniques, the development of faster computers and novel software, and the rivalry of two teams, sped up the process. The human genome was done in 2001. Now, 12 years after the publication of the first fungal genome, over 70 species of fungi have been completely sequenced—including several strains for quite a few species—and many more are in the pipeline. The choice of species to be sequenced was determined by several factors: those that cause human disease or considerable

damage to crops were first, followed by some model organisms like baker's yeast, *Neurospora crassa*, and *Coprinopsis cinerea* (though the data for the ink cap are not yet publicly available). Now, costs have gone down considerably, and the time has come when almost anyone can dream of sequencing his or her favorite fungus (or even him or herself) starting at \$15,000 for material costs. For this amount of money you'll get very raw data produced in a week or so. Mainstream sequencing, analyzing, and annotating of the data still adds up to around \$400,000. Work is scheduled for the false truffle, *Rhizopogon salebrosus* and the dyer's puffball, *Pisolithus microcarpus*, and is well underway for the button mushroom, *Agaricus bisporus*.

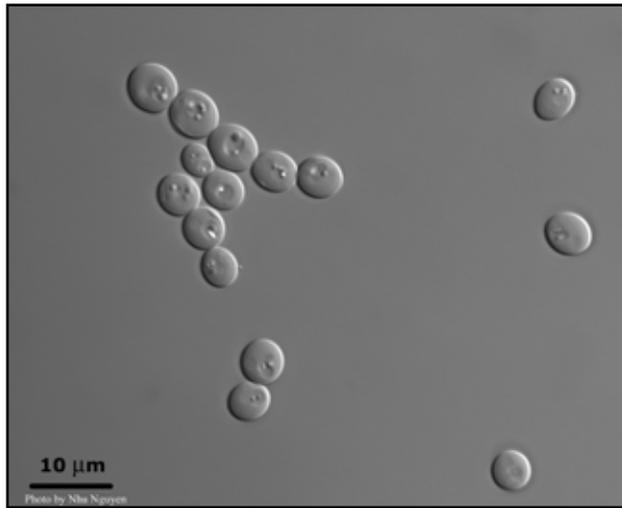


Fig. 2. Yeast cells. Baker's yeast was the first fungus, and the first eukaryote, to be sequenced. Photo by Nhu Nguyen.

Ascomycetes were the first fungi to be sequenced, as that group harbors many well known human pathogens, including *Aspergillus fumigatus*, *Candida albicans*, and *Coccidioides immitis*. Only a small number of basidiomycetes have been sequenced so far. The crust-forming fungus, *Phanerochaete chrysosporium*, was the first basidiomycete. This might be an unknown for the mushroomer, but it is a species with great industrial potential, both as a decomposer of lignin (that hard component of plants, trees in particular)—making it very useful in the paper and fabric industry—and in hazardous waste remediation.

Number two was a human pathogen, *Cryptococcus neoformans*, another yeast. The disease it causes used to be rather obscure, but patients with AIDS, whose immune systems have been compromised are susceptible to this pathogen.

The corn smut, *Ustilago maydis*, and most recently an ectomycorrhizal species, *Laccaria bicolor*, have also been completely sequenced. For all of these, publications can easily be found (see the list under Further Reading), and the data are publicly available on the web. So, for a very varied but extremely limited group of basidiomycetes, the genome data are available.

At first the focus was to figure out what kind of genes there are and what they do. This is certainly a work in progress, as baker's yeast has 6,000 genes, and *Laccaria* around 20,000! Besides genes, the rest of the genomes contain lots of non-coding regions, repeated elements, and so-called junk DNA (all of which has to be sorted out, as well). These pieces can still be useful, though we do not know exactly why and how (that “junk” label may be premature).

The next step was to compare the genes of one species with the genes of others, and relate the differences to lifestyle. The four basidiomycetes, which represent totally different lifestyles, provide a good example. For instance, *Phanerochaete chrysosporium* has many genes involved in the breakdown of lignin, but these genes are lacking in *Laccaria bicolor*. These comparisons also showed that a species often does not have just one gene to perform such an important task, but several, and these might be derived from a single shared ancestor gene. There are also studies looking specifically at those genes that are involved in the decomposition of plant material. These include ones that code for laccases and different types of peroxidases (lignin peroxidases and manganese peroxidases) and indeed, *Laccaria bicolor* has not been equipped with genes for peroxidases. *Cryptococcus* cells are surrounded by a polysaccharide capsule, and this envelope is made by a series of 30 different genes, which are absent in the other basidiomycetes investigated so far. This information might be extremely useful in the battle against this fungus. The smut fungus, *Ustilago maydis*, has, again, a different lifestyle. In one stage it lives as a saprotrophic yeast, and in another grows inside a corn cob, forming the gall-like “huitlacoche,” an enlarged part of the cob full of smut spores. It is not a very aggressive pathogen and lacks the genes to make the enzymes that degrade the plant cell-wall and give it access to the contents. But its genome sequence did reveal an unsuspected set of small genes that play a role in its virulence. In contrast, the rice blast fungus, *Magnaporthe grisea*, an ascomycete, is well-provided with genes that encode for cutinases, the enzymes that decompose cutin (the first barrier the plant uses to keep intruders at bay). Comparing the genetic composition of phylogenetically different fungi with similar lifestyles (e.g. the ectomycorrhizal *Tuber*, an ascomycete, and its basidiomycete counterparts, such as *Boletus edulis* and *Amanita muscaria*), is another interesting research field.

Evolutionary histories of species can be determined by comparing complete genomes, but the small number of fungal genomes available means that these studies still have limited power. One such study, which was published a few years ago, was based on 42 different genomes, of which only four represented basidiomycetes. It would be great if whole-genome studies could indicate which single-gene sequences gave the same results as the more reliable genome-wide phylogenies, in order to validate which sequences to use in future phylogenetic studies. Gene phylogenies are not by definition the same as species phylogenies, as depending on the environmental pressure, genes undergo different changes. The current favorites are LSU and a few protein-coding genes for phylogenetic studies, and ITS as a fungal “barcoder.”

Whole genomes can also reveal aspects of evolutionary history that no single gene can. For instance, they reveal where and when genome duplication took place (as happened once in a group of ascomycete yeasts, close to the baker’s yeast), and they

also show that a switch in the interpretation of the code of the base sequence “CTG” has occurred—in most species this translates into the amino acid leucine, but a group of *Candida* species makes serine out of it.

But up to now, only the surface has been scratched. Coming are more in-depth questions concerning gene function. Does a gene work on its own? When is it active? Does it always have the same function, or does it depend on the circumstances? And of course, many more whole genomes will be sequenced. I’m looking forward to seeing the secrets of my own pet fungi, the beautiful parasol mushrooms, revealed! ☘

#### Further Reading:

- Dujon, B. et al., 2004. Genome evolution in yeasts. *Nature* 430: 35-44.
- Espagne, E., et al., 2008. The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome Biology* 9: R77. [22 pages; open access at <http://genomebiology.com>]
- Fitzpatrick, D.A., M.E. Logue, J.E. Stajich & G. Butler, 2006. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evolutionary Biology* 6: 99. [15 pages; open access at [www.biomedcentral.com](http://www.biomedcentral.com)]
- International Human Genome Sequencing Consortium, 2001. *Initial Sequencing and Analysis of the Human Genome*. *Nature* 409: 860-921.
- IHGSC, 2004. Finishing the *Euchromatic Sequence of the Human Genome*. *Nature* 431: 931-945.
- Kämper, J. et al., 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* 444: 97-101.
- Loftus, B.J. et al., 2005. The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* 307: 1321-1324.
- Martin, F. et al., 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452: 88-93.
- Martinez, D. et al., 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nature Biotechnology* 22: 695-700.
- Morgenstern, I., S. Klopman & D.S. Hibbett, 2008. Molecular evolution and diversity of lignin degrading heme peroxidases in the Agaricomycetes. *J. Mol. Evol.* 66: 243-257.
- Venter, J.C. et al., 2001. *The Sequence of the Human Genome*. *Science* 291: 1304-1351.

<http://fungalgenomes.org/blog/>

[http://fungalgenomes.org/wiki/Fungal\\_Genome\\_Links](http://fungalgenomes.org/wiki/Fungal_Genome_Links)

<http://www.jgi.doe.gov/sequencing/statusreporter/psr.php?projectid=99190> for updates on the *Agaricus bisporus* sequencing project