J. Eukaryot. Microbiol., 51(1), 2004 pp. 30–37


# Comparative Genomics of *Pneumocystis carinii* with Other Protists: Implications for Life Style[1]

MELANIE T. CUSHION

*University of Cincinnati College of Medicine, Division of Infectious Diseases, Cincinnati, Ohio 45267–0560, USA*

ABSTRACT. Three protistan genomes were analyzed for differential genetic traits that may be associated with biological adaptations to their unique life styles. The microsporidian, *Encephalitozoon cuniculi*, an obligate intracellular parasite; the ascomycetes, *Pneumocystis carinii*, considered an opportunistic pathogen; and *Saccharomyces cerevisiae*, a model organism exhibiting a free-living life style, were used in comparisons of genomic architecture, reproductive strategies, and metabolic capacity predicted by the presence of signature genes. Genome size, gene number, and metabolic function decreased as the organisms became more dependent on their hosts. In contrast, gene density and the percentage of genes dedicated to cell growth and division were substantially increased in the genome of *E. cuniculi*. The obligate life style was associated with reductions in gene number, genome size, and reduced metabolic capacity while the free-living life style was coincident with gene duplications and duplication of large portions of the genome. The genomic characteristics and metabolic capacity of *P. carinii* were usually intermediate between those of the other two protistan genomes, but unique characteristics such as the presence of a single rDNA locus may indicate that these organisms could be in the process of becoming more host dependent.

**Key Words.** Free-living, model organism, obligate intracellular parasite, opportunistic pathogen.

SEQUENCING of whole microbial genomes began with the bacterial pathogen, *Haemophilus influenzae*, in 1995 (Fleischmann et al. 1995). Release of the first eukaryotic genome sequence, the model organism, *Saccharomyces cerevisiae*, soon followed in April of 1996 (Goffeau et al. 1996), with a complete directory published in 1997 (Goffeau 1997). To date, 362 other prokaryotic genomes and 243 eukaryotic genomes plus 8 chromosomes have been sequenced or are in the process (Integrated Genomics 2003). Of the 152 published genome sequences, the vast majority are prokaryotes (78%) while only 12% are eukaryotic, and of those, only about a quarter are considered "protists". This situation will soon reach a better balance with the completion of those projects that are ongoing, especially due to the efforts of the Fungal Genome Initiative spearheaded by the Whitehead Institute (Whitehead Institute 2003). Fifteen fungal genomes are currently under construction, while another 44 have been proposed for future projects.

Rationale for the selection of protistan genomes to be sequenced can be placed in these general categories. (1) Model organisms. Fungi such as *S. cerevisiae* and *Neurospora crassa* have the advantages of a large community of investigators that can benefit from the whole genome sequence, while providing annotation capacity by virtue of a large sink of previously sequenced genes and DNA regions. These organisms often have well-developed experimental systems that provide the ability to identify genes of unknown function by deletion, mutation or similar analyses. (2) Pathogens. Pathogenic protists are often not as tractable as the model organisms. However, their detrimental effects on human health, the domestic farm animal and agricultural industries, combined with the promise of new treatments or control measures as a result of the whole genome sequence, give organisms like *Plasmodium falciparum* and even its vector, *Anopholes*, high priority status. (3) Organisms of evolutionary interest. Phylogenetic placement and evolutionary concepts have undergone dramatic changes within the last 2 decades due to the routine availability of sophisticated DNA sequencing reagents and instrumentation, and the application of mathematical concepts for evaluation of evolutionary time as a measure of change in DNA sequence. Comparative genomics of organisms from different kingdoms and branches as well as from within clades facilitate understanding of the evolution of pathogenic processes, genes and their regulation, biochemical pathways, and chromosomal architecture. Comparative genomics can also provide insights into strategies that organisms have evolved for adaptation to specific habitats and niches. In the present analysis, 3 protistan genomes were evaluated for potential differences that may be related to their very distinct life styles.

**Selection of protistan genomes.** The genomes of *Encephalitozoon cuniculi*, *Pneumocystis carinii*, and *Saccharomyces cerevisiae* were chosen for this analysis because of their specific life styles and availability of sequence. All 3 organisms are considered members of the Fungal Kingdom, with *E. cuniculi* being the most recent addition (Hirt et al. 1999). *Encephalitozoon cuniculi* is considered an obligate intracellular parasite; *P. carinii* has been designated as an opportunistic pathogen; and *S. cerevisiae* represents a free-living protist.

## IDENTIFICATION AND TAXONOMY

***Encephalitozoon cuniculi.*** The first microsporidian identified was the agent of "pebrine disease" which almost destroyed the silk industry in Europe by causing a fatal infection of the silkworms. In 1857, Näegli found the microscopic agent causing the disease, which he thought had morphological similarities to yeast and called it "*Nosema bombycis*", placing it in the "schizomycete fungi" a group that included an assortment of yeasts and bacteria (Keeling and Fast 2002). The agent of the first microsporidial infection reported in mammals was reported in 1922 and named *Nosema cuniculi*, which caused an infectious paralysis in young rabbits (Wright and Craighead 1922). Levaditi, Nicolau, and Schoen assigned this new species a new genus, *Encephalitozoon* in 1923 (Weiser 1985). It is a member of a phylum that contains over 150 genera and 1,200 species (Keeling and Fast 2002).

In 1882, Balbiani created a new phylogenetic group for the *Nosema* species, Microsporidia, which until recently, was where it resided. The Microsporidia were placed in another "grab bag" of protozoans called the Sporozoa in 1947 due to their ability to form spores (Kudo 1947). The Sporozoa as a taxonomic term was later discredited and members dispersed to apicomplexa, myxosporidia, haplosporidia, and several other groups. A clear understanding of this group of microbes remained elusive. In 1983, Cavalier-Smith proposed the formation of the Archezoa, a group of primitive eukaryotes that he hypothesized had evolved prior to the endosymbiotic origin of mitochondria (Cavalier-Smith 1983). Besides the Microsporidia, 3 other amitochondriate lineages were included; Archae-

Corresponding Author: M. T. Cushion—Telephone number: 513-861-3100 ext. 4417; FAX number: 513-475-6415; E-mail: Melanie.cushion@uc.edu or E-mail: Melanie.Cushion@med.va.gov

[1] Symposium presentation "Advances in Protistology," June 15, 2003, during a joint meeting of the Society of Protozoologists and Phycological Society of America, Gleneden Beach, Oregon, USA.

moebae, Metamonada, and Parabasalia. In addition to the lack of mitochondria, other "primitive" characteristics were identified in or lacking from the genomes of the microsporidia as evidence for their early lineage. These included a 70S ribosome rather than the larger 80S ribosomes of other eukaryotes; the absence of flagella and other similar structures with the 9+2 organization; and a fused 5.8s RNA/large subunit rRNA (Vossbrinck and Woese 1986). The latter characteristic had only previously been found in prokaryotes. Phylogenetic analyses based on the small subunit rRNA, the molecular chronometer of choice, showed the microsporidian, *Vairimorpha*, to occupy the earliest branch of the eukaryotic lineage (Vossbrinck et al. 1987).

With increased genetic data available, the primitive nature of the Microsporidia was refuted and a new concept of these organisms as "highly evolved fungi" was put forth in 1998 (Keeling and Fast 2002; Keeling and McFadden 1998). Although the sequence analyses previously supported the contention of Cavalier-Smith and others, Keeling noted that the highly divergent microsporidial gene sequences could have resulted in misleading phylogenetic trees. Moreover, the sequence of the fused 5.8SLSU rRNA was not only divergent from other eukaryotes, but prokaryotes as well, and the fusion may have arisen secondarily as a result of deletions that reduced the size of the molecule (Cavalier-Smith 1993). In contrast to the earlier gene phylogenies, subsequent analyses with genes such as alpha- and beta tubulin, TATA-box binding protein, mitochondrial Heat Shock protein 70, glutamyl tRNA synthetase, and the large subunit of the RNA polymerase II, showed a clear alignment with the fungi (Keeling and Fast 2002). The apparent "amitochondriate" state of the microsporidia is currently being questioned due to the presence of mitochondrial-specific genes in the genomes of some microsporidia, e.g. HSP70 (Hirt et al. 1997). Indeed, sequencing of the *E. cuniculi* genome permitted identification of a set of genes that were of mitochondrial origin, prompting investigators to propose the existence of a cryptic organelle, the "mitosome" in *E. cuniculi* (Katinka et al. 2001; Vivares et al. 2002b). This finding is of critical importance to our understanding of the phylogenetic status of the microsporidia. All fungi to date contain mitochondria and with the emergence of evidence that implies the microsporidia did contain these organelles at one point in their ancestry, the concept of microsporidia as primitive organisms has now evolved to one that embraces them as "highly evolved fungi"(Keeling and Fast 2002). Some organisms adapt to their environments by processes such as gene duplication or uptake of genetic elements from foreign genomes. In the case of microsporidia, a "reductive evolution" appears to be the adaptive mechanism it used for life as an intracellular parasite. Their small genome size, reduced gene lengths, and gene compaction (discussed below) support this contention.

***Pneumocystis carinii.*** Like *E. cuniculi*, the classification and taxonomy of the organisms in the genus *Pneumocystis*, has not been a straightforward process. In 1909, Carlos Chagas, a noted Brazilian researcher, was studying the life cycle of American trypanosomes. Bringing together his microscopic observations in humans, rodents, guinea pigs, and non-human primates he proposed an elaborate life cycle for these parasites that included a sexual cycle that took place in the lungs (Chagas 1909). Through an intermediary, Antonio Carini, who was studying at the Oswaldo Cruz Memorial Institute at that time, the husband and wife research team of Msr. and Mme. Delanöe received slides of the lung parasites a few years later, reviewed in Hughes (1987). They felt the lung organisms were not a part of the trypanosomal life cycle, but a completely different genus and species. A series of experiments that were elegant for the time, involving isolation of rats, dams, and pups, showed that rats without trypanosomal infection were infected with this new organism. They called it *Pneumocystis carinii*, to denote its predilection for the lung and its characteristic developmental form, the cyst, and to honor Dr. Carini for his role in the process. Like Chagas, the Delanöes thought the organisms were protozoan parasites, but the question of whether they were fungi was raised throughout the ensuing decades. Prior to the use of electron microscopy, morphological studies like those of Giese in 1953 likened the morphology of the cysts to the asci in yeast and were compelled to argue for its fungal nature (rev. Cushion 1998a). Ultrastructural studies from the late 1960s and 1970s reviewed the morphological affinities of the developmental forms of rat and human *Pneumocystis* in the respective lungs and reached the opposite assignments to Fungi (Vavra and Kucera 1970) and to Protozoa (Barton and Campbell 1969; Barton and Campbell 1967; Campbell 1972), underscoring the dilemma in the field that persisted up until the end of the 20th century. It wasn't until the late 1980s that molecular genetic techniques were applied to analysis of the genes of *P. carinii* for phylogenetic determinations. Two laboratories co-incidentally sequenced the nuclear small subunit rRNA as rRNA (Stringer et al. 1989a; Stringer et al. 1989b) and rDNA (Edman et al. 1988) and both reached the same conclusion that *P. carinii* was a member of the Fungi. Since then, all genes that have been analyzed for phylogenetic assignment have supported the fungal nature of *P. carinii*. Significantly, the *P. carinii* sequence derived from the *Pneumocystis* Genome Project (Cushion and Arnold 1997) showed most of the putative genes to be most similar to genes from members of the Fungi (Cushion and Smulian 2001), specifically the fission yeast *Schizosaccharomyces pombe*. It now seems clear that *Pneumocystis* is a member of the Fungi. The closest extant relative appears to be *S. pombe*, but additional contributions to the gene databases may eclipse this current association. Recognizing the fungal nature of *Pneumocystis*, the following classifications have been proposed: Ascomycota; Pneumocystidomycetes; Pneumocystidales; Pneumocystidaceae; Pneumocystis (Erikkson 1994; Taylor et al. 1994).

***Saccharomyces cerevisiae*** **Meyen ex E.C. Hensen 1883.** There has been little doubt as to the fungal nature of the budding yeast, *S. cerevisiae*. Probably the most notable aspect is the multitude of synonyms (> 100) associated with the name based on apparent phenotypic differences (Barnett et al. 1990). Its classification is as follows: Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

## LIFE STYLES AND LIFE CYCLES

***Encephalitozoon cuniculi.*** This is a small (1.5 × 2.5 mm) organism that maintains an obligate intracellular parasitic existence in mammals, where it can infect a wide variety of hosts including insectivores, rodents, carnivores, and primates (Vivares and Metenier 2001). This species was the first microsporidian to be propagated in cell culture (Vivares and Metenier 2001). *Encephalitozoon cuniculi* has mainly been associated with infections of the eye, respiratory and genitourinary tracts, and disseminated infection (Centers for Disease Control and Prevention and Division of Parasitic Diseases 2003). The spores of microsporidia are the only stage that can exist outside the host cell. Spores are found in a wide variety of environmental niches, including ditch water and can infect the host through almost any route. The site of infection is somewhat dependent on the route of transmission. Inhaled spores usually lodge in the lung, causing respiratory complications while ingested spores parasitize the cells of the gastrointestinal tract. However, these organisms can clearly travel to other sites since a myriad

of infections have been attributed to microsporidia, including chronic diarrhea and wasting; keratoconjunctivitis; pneumonia; bronchitis; nephritis; urethritis; prostatitis; hepatitis; encephalitis; myosititis and peritonitis.

The most striking life-cycle stage of any of the microsporidia is the spore. The morphological characteristics of each spore aid in differentiation of the species. Spores range in size from 1- to 40 μm and can be ovoid, spherical, rod-shaped or even crescent-shaped. Their walls contain alpha-chitin which contributes to survival in the external environment. The most obvious component of the spore is the polar filament (tube) which is a highly specialized super-coiled structure that is connected to an anchoring disk at the anterior and terminates at the posterior vacuole. The coil formation is characteristic within each species and ranges in length from 50- to 500 μm in length and from 0.1- to 0.2 μm in width. In *E. cuniculi* spores, there are 5–7 coils in a row. Upon germination, which is initiated by factors that are poorly understood, but certainly upon contact with the target host cell, the osmotic pressure increases within the spore to a point that causes a rapid ejection of the polar filament. As the polar filament breaks through the spore and into the target cell, the entire contents of the spore including the nucleus and the spore contents (sporoplasm), turn inside out and are forced through the filament and injected into the host cell. This eversion event takes place at the incredible speed of 100 μm/sec for a sum total of 2 sec to complete the process (Frixione et al. 1992). The parasitic stage inside the host cell, the meront, begins successive generations of asexual growth via binary fission (merogony) within a parasitophorous vacuole in the host cytoplasm. The use of the vacuole is in contrast to other microsporidians that are in direct contact with the host cytoplasm. The parasitophorous vacuole is surrounded by host mitochondria, which suggests this recruitment to be a strategy by the amitochondriate parasite to secure an energy source. Merogony is followed by sporogony, which includes production of sporonts that have a thickened coat. The sporonts divide into sporoblasts, followed by maturation to spores.

***Pneumocystis carinii.*** This organism is considered an "opportunistic pathogen" but the accuracy of this designation has been disputed (Frenkel 1999). Other organisms considered to be opportunistic pathogens, such as *Aspergillus fumigatus*, have clear environmental habitats such as in rotting organic waste, but are able to grow and cause disease in mammals by virtue of their ability to survive at higher temperatures and other potential virulence factors that provide protection in the hostile mammalian environment. To date, no environmental reservoir has been identified for any member of the *Pneumocystis* genus and increasing evidence suggests that this family of organisms can reside in their specific immune-competent host, perhaps as a cryptic infection, with little clinical manifestation (Icenhour et al. 2001; Icenhour et al. 2002; Vargas et al. 2000; Vargas et al. 2001). Unlike *E. cuniculi* or *S. cerevisiae*, no species of *Pneumocystis* has been propagated continuously outside the mammalian lung (Cushion 1989a).

*Pneumocystis* spp. cause a pneumonia which can become lethal if the host enters an immunosuppressed state. Thus it has been suggested that the infection be referred to as an immunodeficiency-dependent disease (IDD) (Frenkel 1999). Unlike *E. cuniculi*, members in the genus *Pneumocystis* exhibit strict host specificity with at least 1 distinct species resident in the lungs of every mammalian host examined for its presence (Cushion 1998b). Also in contrast to the microporidial organism, all species of *Pneumocystis* prefer the lungs as their primary niche within the mammalian host. Reports of disseminated pneumocystosis are relatively infrequent, but when these organisms venture out of the lungs, they infect a wide variety of cells and organism systems, with some preference for the bone marrow and spleen, and an avoidance of brain tissue (Cote et al. 1990; Telzak et al. 1990).

The complete life cycle for any *Pneumocystis* species is not known. However, ultrastructural and light microscopic observations of the organisms within the lung have led to several hypothetical cycles, reviewed in Cushion (1998a). There is however, general agreement as to the likely existence of at least 2 phases of the life cycle that occur within the mammalian alveolus. In the asexual phase, the smaller trophic forms (1–4 μm), replicate by binary fission, like merogony in *E. cuniculi*, but in the extra-cellular milieu of the alveolus and not within a host cell. Evidence for a sexual stage was first reported in an ultrastructural study which identified synaptonemal complexes in a stage the authors referred to as an "early precyst" (Matsumoto and Yoshida 1984). Recent molecular genetic data support the presence of meiosis in *P. carinii*. A Ste3-like pheromone receptor was identified in an EST database and subsequently, in a cosmid clone from the *P. carinii* genome (Smulian et al. 2001). In yeast, Ste3 functions in concert with the a-factor from the opposite mating type (which has a Ste2 receptor for the alpha pheromone) to initiate mating, resulting in meiosis. This process is mediated by a signal transduction cascade of which some putative gene homologs in *P. carinii* have been identified as well (Smulian et al. 2001; Vohra et al. 2003). In addition, homologs to meiosis-specific genes in fungi (e.g. *Mei1*) have been identified by the Pneumocystis Genome Project (Slaven et al. 2003). The sexual cycle of *P. carinii* likely results from the conjugation of haploid trophic forms of opposite mating type that proceeds to form a diploid zygote resulting in 4 haploid nuclei in a form called the "precyst" (Cushion 2003). An additional mitotic event results in 8 haploid nuclei that are then compartmentalized within the maturing cyst to produce spores within the spherical and rigid cell wall. Release of the spores may involve escape through a rent in the cell wall, but this process has not been adequately documented. It is presumed that once the daughter forms are released, they become the vegetative trophic forms. The infectious propagule of *Pneumocystis* spp. has not been identified, but the infection has been shown to be transmitted by an airborne route (Hughes et al. 1983; Walzer et al. 1977), but could also involve close contact (Icenhour et al. 2002). The route of escape from the host by the agent of infection is also unknown.

Both the asexual and sexual cycles appear to occur within the mammalian lung and in close juxtaposition. Attachment to the Type I alveolar cell by the trophic form may be the initiating factor for the infection. The trophic form and the host cell are in close apposition to each other, but the organisms remain extra-cellular. As the infection progresses, large clusters of tightly adherent trophic forms, intermediate cyst stages, and cysts co-exist in the lumen of the alveoli, filling it until the exchange of gas is compromised and the host enters respiratory failure.

***Saccharomyces cerevisiae.*** The natural habitats of *S. cerevisiae* are those where it is moist with an abundant supply of simple, soluble nutrients such as sugars and amino acids. Common substrates include leaf and fruit surfaces, roots, in various types of food and on mammalian skin. Like *P. carinii*, and unlike *E. cuniculi*, *S. cereviaie* has both an asexual and sexual cycle. Simple budding of diploid cells is the mode of replication when nutrients are plentiful. If starved, they undergo meiosis giving rise to haploid spores. These spores can then germinate under improved conditions and asexually proliferate or undergo conjugation (mating) to reconstitute the diploid phase. Budding is initiated during $G_1$ and as the mother cell increases in size, the nucleus is replicated and spindles are formed during the S

Table 1. Unique characteristics.

| *Encephalitozoon cuniculi* | *Pneumocystis carinii* | *Saccharomyces cerevisiae* |
|---|---|---|
| Obligate intracellular life style | "Opportunistic" extracellular life style | Free-living life style |
| No ergosterol | No ergosterol | Ergosterol is bulk sterol |
| No mitochondria | Typical fungal mitochondria | Mitochondria |
| Subtelomeric arrays of rDNA: 1 locus per chromosome end (22 total) | Single copy of rDNA locus | ~150 copies of rDNA locus: tandem array on chromosome XII |
| Single membrane in sporoplasm; double in sporoblast; triple in spore | Double membrane throughout the life cycle | Single plasma membrane |
| Chitinous cell wall in spore | Delicate cell wall in precyst to cyst stage | Rigid cell wall |
| Cultivation on mammalian cells | Intractable to long term culture | Grows easily on artificial medium |
| Asexual cycle; binary fission | Asexual cycle; binary fission | Asexual cycle; budding |
| No sexual cycle | Sexual cycle | Sexual cycle |

phase, the mitotic entry checkpoint occurs after the S phase followed by establishment of the septum and the pinching off of the daughter cell. The age of a cell can often be determined by the number of bud scars formed after each replication.

In order to fuse, the haploid cells of yeast must be of opposite mating types, either alpha or a. Each produces a mating factor (pheromone) that is complementary to the receptor (Ste 2 or Ste 3) on the opposite mating type. The receptor-pheromone system facilitates cell contact and fusion, initiating the mating process which is mediated by signal transduction cascades. The meiotic process results in an ascus containing 4 haploid spores. Determination of the yeast cell mating type is controlled by 3 genes in the *MAT* locus, *alpha 1, alpha 2* in the alpha mating type and *a1* in the a type. Once the two types have mated, a unique pattern of gene expression occurs distinct from either mating type. The *S. cerevisiae* genome contains copies of both mating type regulatory genes on either side of the *MAT* locus. Mating types can switch after every other cell division using a gene conversion process called the "cassette mechanism", which involves the excision of an entire copy of the active cassette and replacement by a newly synthesized copy of the silent cassette of the opposite mating type (Herskowitz 1989).

In addition to budding and mating, diploid *S. cerevisiae* strains are able to undergo a morphological change coincident with a polar cell division resulting in invasive filamentous growth (Gimeno et al. 1992). The yeast cells elongate and form long chains, called "pseudohyphae" in response to starvation for nitrogen. As the pseudohyphae grow in a direction away from the colony, they invade the agar medium, rather than remaining on its surface. It has been surmised that this adaptation permits yeast cells in their natural habitat to forage for nutrients.

These 3 fungi differ dramatically in their life styles and substrate (host) requirements. *P. carinii* and *S. cerevisiae*, unlike *E. cuniculi*, undergo asexual as well as sexual reproductive phases, but differ in the process, using budding or binary fission. It is yet unclear whether *P. carinii* possesses a mating type cassette system reminiscent of *S. cerevisiae*. Although *E. cuniculi* has dispensed with the process of sex, it produces environmentally resistant spores, which *P. carinii* and *S. cerevisiae* also appear to do.

Each of the organisms has adapted to their specific host milieu. *Encephalitozoon cuniculi* has a complex method of invading target host cells and an efficient method of replication within the host cell shielded within a parasitophorous vacuole. Like *P. carinii*, *E. cuniculi* lacks ergosterol, which is the bulk sterol of most other fungi, and may use host cholesterol as its primary sterol. It also lacks a fully functioning mitochondrion, likely drawing energy from the host's mitochondria. Like *P. carinii*, it takes advantage of multiple membranes to shield itself from the environment—with a single membrane in the sporoplasm, a double in the sporoblast, and a triple in the spore.

Although the entire life cycle of *P. carinii* is not known, other studies showing the ubiquitous nature of the organism in healthy rat colonies (Icenhour et al. 2001), suggest a highly efficient method of transmission. Its ability to survive in the hostile environment of the lung is a topic of ongoing research. It may take up host cholesterol, or it may synthesize it, as suggested by the presence of several key enzymes in the sterol biosynthesis pathway in its genome (Slaven et al. 2003). *Pneumocystis carinii* has a fully functioning mitochondrion with typical fungal elements (Cushion et al., in press 2004) which is the target of drug therapy, e.g. atovaquone. One of the most unique characteristics of the *P. carinii* genome is the presence of only a single nuclear ribosomal locus (Giuntoli et al. 1994), which differentiates it not only from other fungi that often have over one hundred copies, but from most eukaryotes in general. The purpose of this adaptation or the methods for maintaining its genetic stability are not known.

*Saccharomyces cerevisiae* has evolved to maintain reproductive capacity in times of abundance and during starvation. The production of pseudohyphae suggests that it can even seek out sources of nutrients.

A summary of the unique characteristics of these organisms is listed in Table 1.

## INSIGHTS INTO ADAPTATION AND LIFE STYLES BY GENOMIC ANALYSES

The sequence of the genome of *S. cerevisiae* was reported in 1996 as a result of the efforts of a community of over 100 international investigators (Goffeau et al. 1996) and later as a supplement in 1997 (Goffeau 1997). In 2001, the genome sequence of the microsporidia, *E. cuniculi*, was published by investigators at Genescope, Universite Blaise Pascal and the Université Lyon in France (Katinka et al. 2001) with additional analyses of the genome at a later date (Vivares et al. 2002b). The sequencing of the genome of *P. carinii* is ongoing with frequent updates posted on the website (Slaven et al. 2003). A partial EST database is being collapsed to a unigene set. BLAST analysis and assemblies are available on the same website. Reports on the progress of the project are summarized in the following references: (Cushion and Smulian 2001; Cushion et al. in press, 2004; Keely et al. 2001; Keely et al. in press, 2004). Although the project is not yet complete, there are clear indications of the metabolic capacity of *P. carinii*. It should be noted that the comparisons of the genomes discussed herein are limited not only by the limited data set from *P. carinii*, but also from *E. cuniculi* with only about 50% of the potential genes currently identified. Even the genome of the model organism,

Table 2. Genomic organization.

| Genomic feature | *S. cerevisiae* | *P. carinii* | *E. cuniculi* |
|---|---|---|---|
| Size (Mb) | 13.4 | 8 | 2.9 |
| No. genes | ~6,000 | ~4,000 | ~2,000 |
| Av. intergenic distance (bp) | 630–945 | 498 | 120 |
| Gene density/chromosome core | 1 CDS/2000 bp | 1 CDS/2139 bp | 1 CDS/1025 bp |
| Av. no. introns | ~250 total | 3.7/Gene | 15 total |
| Av. intron size range (bp) | 30–40 | 31–235 | 23–52 |
| No. rDNA loci | ~150 | 1 | 22 |

*S. cerevisiae*, has an estimated one-sixth of its genes still classified as "orphans" (Herrero et al. 2003). It is recognized that with improved gene finding programs and other software to place genes into metabolic cycles, the analyses here may require modification. With these provisos in mind, there are still many characteristics of the genomes relating to life styles that can be drawn.

**Genomic organization.** Table 2 provides a summary of the characteristics of the 3 genomes. One of the most dramatic differences among the genomes is the decrease in sizes as the organisms change to a more obligate life style. The free-living *S. cerevisiae* has a genome size of about 13 Mb distributed into 16 linear chromosomes that range in size from 2,200 kb to 200 kb. The genome of the opportunistic pathogen, *P. carinii*, is about 8 Mb and distributed into about 17 linear chromosome-sized bands ranging in size from 700 kb to 300 kb, some of which appear to co-migrate on pulsed field electrophoresis gels (Cushion et al. 1993). The genome of *E. cuniculi*, the obligate parasite, is a mere 2.9 Mb that is nonetheless separated into 11 chromosomes that migrate between 315 to 217 kb (Vivares et al. 2002a). It has the smallest known eukaryotic genome.

This pattern of reduction is repeated in the number of genes per genome and average intergenic distance (Table 2). In contrast, the gene densities between *S. cerevisiae* and *P. carinii* are similar, at about 1 gene per 2,000 bp. However, this distance is halved in the *E. cuniculi* genome. *Pneumocystis carinii* contains the highest number of introns with an average of about 3.7 per gene, while in the entire yeast genome there are very few, about 250 total, with even less in the microsporidia's genome, 15 total. *P. carinii* also has the largest-sized introns, although most are within the size range of typical yeast introns (Thomas et al. 1999). The *P. carinii* genome also stands out because of its paucity of rDNA genes, with its single locus.

The genomic strategies of the 3 organisms seem to reflect their life styles. *S. cerevisiae*, seeking a simplified nutrient substrate on fruits and other organic sources, must withstand temperature and other environmental pressures for survival. It has devised a reproductive scheme to take advantage of times of plenty (asexual replication) and nutrient deprivation (sexual reproduction and spore formation). It is widely held that the haploid genome of *S. cerevisiae* arose from an evolutionarily ancient polyploidization event, due to the presence of numerous duplicate blocks in the present day genome (Hughes and Friedman 2003; Wolfe and Shields 1997). In addition, the *S. cerevisiae* genome has gene duplications that took place independently of the polypolidization (Hughes and Friedman 2003). In a genomic analysis of *S. pombe* and *S. cerevisiae*, duplication in both genomes was observed in 56 gene families that were mostly associated with cell growth, fission, and ribosomal proteins, leading the authors to suggest that at least these duplications may be involved in the adaptation to a unicellular yeast life style (Hughes and Friedman 2003). The expansion of the *S. cerevisiae* genome provides it with abundant reproductive capacity and strategies for survival.

At about half the genome size and 66% of the gene complement of *S. cerevisiae*, *P. carinii* shows a reduced capacity in its life style as compared to that of the yeast. There is no evidence that *P. carinii* has an environmental cycle, but there is increasing evidence indicating the mammalian host is its sole niche. Within its genome, the spacing of the genes is closer than that in the yeast, suggesting conservation of the genome. However, the genes of *P. carinii* are studded with introns, sometimes up to 9, as in the *pcg1* gene (Smulian et al. 1996). The presence of introns could be considered costly in terms of energy and machinery necessary for processing. However, the organism may be using splice intermediates as a strategy to increase its genetic repertoire as in the case of the gene encoding the inosine monophosphate dehydrogenase protein. In *P. carinii*, the IMP dehydrongenase gene with a total of 5 introns was found to exhibit 3 splice variants that were preferentially expressed in vivo or in vitro (Ye et al. 2001). The variant that was expressed in vivo and was catalytically active contained introns 1, 3 and 5, suggesting that the organism may use differential splicing as a means to respond to different environmental pressures.

Another way in which *P. carinii* may fully utilize its genome is by exploiting both DNA strands to encode genes. Sequence from a 32-kb cosmid clone revealed the presence of 6 genes on the Watson strand and 9 genes were found on the Crick strand (Smulian et al. 2001)

The presence of a single copy of the rDNA locus in the genome of *P. carinii* contrasts with the latest studies that show a correlation between rDNA copy number and genome size in eukaryotes (Prokopowich et al. 2002). The presence of a single rDNA locus may at first appear counter-intuitive to survival of the cell. Once the gene function is lost, due to mutation or drug therapy, for example, the viability of the organism ceases. The inability of the organism to grow in artificial medium or on host cells outside the mammalian lung could possibly be a consequence of this phenomenon, yet ultrastructural studies indicate that ribosomes are plentiful and growth within the mammalian lung, if unchecked, can fill the alveoli with millions of parasites. However, the rate of growth in mammalian lung appears to be slow to reach the point of fulminate infection. In the rat model of *P. carinii* pneumonia, approximately 8–12 wk of chronic immunosuppression is required to reach peak infection, whether organisms are inoculated directly into the lung via the trachea, or if the animals obtain the organisms by the natural route of exposure. Perhaps the slower rate of growth aids the organism in maintaining low numbers within the immune competent host, but once the host becomes more hospitable by loss of immune function, the organisms respond to the more permissive environment and increase their rate of growth. However, deletion of rDNA copies in yeast systems have not equivocally showed a resultant decrease in growth rate (Prokopowich et al. 2002). But, it should be noted that most yeasts contain numbers of rDNA copies well in excess of that necessary to maintain cellular function, and the minimal threshold

Table 3. Percent of genes in functional categories.

| Functional category | *E. cuniculi* | *P. carinii* | *S. cerevisiae* |
|---|---|---|---|
| Metabolism | 8.4 | 6 | 17 |
| Energy | 2.6 | 2 | 3 |
| Cell growth, division, and DNA dynthesis | 15.4 | 7 | 14 |
| Transcription | 19.6 | 9 | 10 |
| Protein synthesis | 14 | 5 | 5 |
| Protein destination | 14.1 | 18 | 7 |
| Transport | 12.6 | 6 | 10 |
| Cellular organization/biogenesis | 7.7 | 12 | 28 |
| Cellular communication/signal transduction | 3.1 | 0.4 | 2 |
| Cell rescue, death and aging | 2.5 | 1 | 4 |

may not have been reached in these studies. Certainly, no study has ever decreased the copy number to one. The enigma of the sole rDNA locus in *P. carinii* awaits further study.

Several ends of *P. carinii* chromosomes have been sequenced and analyzed for structure (Keely et al. 2001). Unique to this organism is the presence of repetitive units of genes encoding surface antigens referred to as "major surface glycoproteins" (MSG gene family), MSG-related proteins of unknown function (MSR gene family) and subtilisin-like proteases (PRT gene family) that may be involved in the processing of the MSG proteins (Stringer and Keely 2001). These genes are arranged in 5′ to 3′ tandem arrays and contain varying numbers of the MSG, MSR, and PRT genes. Downstream of these gene families are a series of subtelomeric repeat regions terminating in telomeric repeats (TTAGGG). The MSG and MSR gene families appear to be attached to the ends of all the chromosomes and are quite large in size, ~ 30 kb. The sum total of DNA dedicated to these genes comprises approximately 10% of the genome (Stringer and Cushion 1998). It is currently held that the vast repertoire of genes encoding surface antigens may permit the organism to somehow evade the host immune system. Another consideration may be a role in its life cycle, such as in mating. Adhesins found in other fungal genomes mediate the close proximity of the 2 mating types (e.g. FLO genes in *S. cerevisiae*). The MSG proteins are clearly adherent and potentially could function in this manner. The investment of the organism in these gene families certainly makes it appealing to speculate that they are a primary survival mechanism.

The genome of *E. cuniculi* is a bit more than one-third the size of that of *P. carinii* and about 25% the size of the yeast genome. It has only one-third of the genes of *S. cerevisiae* and half of the complement of *P. carinii*. Its genome is quite gene dense, with genes following almost one after the other. Concomitantly, the size of the genes are smaller than those in other protists. About 85% of its protein-encoding genes were smaller than yeast homologues, with an average decrease in length of about 15% (Katinka et al. 2001). Together with the relative lack of introns, the 2.9 Mb of *E. cuniculi* contains mostly functional units. The truncation of genes has been postulated to result in decreases of protein-protein interactions (Vivares et al. 2002b), which in effect translates to decreased functions that are presumably replaced by the host. It seems apparent that the reductionist strategy taken by this microsporidian is in direct response to its habitat within the host cell as an obligate organism.

**Functional capacity predicted by genomic analyses.** The genes identified in the *E. cuniculi* (Katinka et al. 2001; Vivares et al. 2002b) and *S. cerevisiae* (Martinsrieder Institut für Protein Sequenzen 2003; Mewes et al. 1997) genomes and the EST database and available genomic sequence from the *P. carinii* genome project (Slaven et al. 2003) were classified for functional activity using the Munich Information Centre for Protein Sequences (MIPS) categories. Table 3 summarizes the data available from each genome. Approximately 44% of the potential genes of *E. cuniculi* did not have any identifiable homologues; 21% of the EST contigs analyzed for *P. carinii* were considered "No Hits" and about 50% of the data set used for *S. cerevisiae* did not have functional identity. More recent functional categorizations have dramatically reduced the number of these unknown ORFs, but reclassification schemes placed genes in more than a single category (Herrero et al. 2003), prohibiting a direct assessment among the 3 genomes being analyzed. The earlier data were used in the present study.

Although diminished in genome size and number, *E. cuniculi* maintains a large contingent of genes whose function it is to replicate the organism (Cell growth, Division and DNA Synthesis, Transcription) (35%) and for protein trafficking and degradation (28%). Genes devoted to transport are higher than in the other protists (~ 13%), which may be a function of its intracellular location and requirement for compounds within the host cytoplasm to be shuttled across its membranes.

The genes identified for metabolism and energy in the *P. carinii* genome thus far are slightly less than those found in the other 2 protists. Likewise, the numbers for Cell Growth, Division, and DNA synthesis, Transport, Cellular Communication and Cell Rescue are also decreased. It is likely this is a function of the incomplete state of the genome rather than the state of its metabolic capacity. Noteworthy are the similarities to *S. cerevisiae* in the percentage of genes contributing to Transcription and Protein Synthesis and the high percentage associated with Protein Destination. In spite of these comparatively lower numbers, BLAST analysis of potential gene homologs infers an active metabolism for *P. carinii* that includes purine and pyrimidine biosynthesis, amino acid synthesis, sterol biosynthesis, TCA cycle and electron transport chain, glycolysis, and trehalose metabolism (Table 4).

## SUMMARY

The obligate life style exemplified by *E. cuniculi* illustrates how few genes may be necessary to ensure survival, even in hostile environments. The microsporidial ancestors likely had a mitochondrion, but lost the genes and now rely on the host for most of their energy requirements. The spore is the only life-cycle stage outside the host and it is extremely resistant permitting its survival in harsh conditions. The infective apparatus within the spore is a highly evolved and sophisticated mechanism, the sole purpose of which is to invade the target cell. Once inside the cell, the organism replicates rapidly, with a large contigent of genes within its minimal genome devoted to this process. The lack of DNA-containing organelles like the mitochondrion may also make replication more efficient. Likewise, required compounds it no longer synthesizes (e.g. amino

Table 4. Notable differences and similarities.

| *E. cuniculi* | *P. carinii* | *S. cerevisiae* |
|---|---|---|
| No purine biosynthesis | Purine biosynthesis | Purine biosynthesis |
| 2 Amino acid synthesis genes | Amino acid synthesis genes | Amino acid biosynthesis |
| No sterol or fatty acid biosynthesis genes | Sterol and fatty acid synthesis genes | Sterol and fatty biosynthesis |
| No TCA cycle | TCA and electron transport | TCA and electron transport |
| "Mitosome" no mitochondrion | Mitochondria | Mitochondria |
| Glycolysis | Glycolysis | Glycolysis |
| Trehalose metabolism | Trehalose metabolism | Trehalose metabolism |

acids) are supplied by the host via a vigorous transport system. Although reduced in size, the genome makes full use of its DNA by very high gene density, gene compaction, and few introns.

The "opportunistic pathogen" *P. carinii* has a larger genome, more genes, and makes use of its DNA by encoding genes on both strands and perhaps by use of alternative splice intermediates. These organisms have a more complete metabolic capacity than the microsporidia, but it is difficult to tell at this time if all metabolic cycles are present. From the available data, it appears that *P. carinii* contains many biosynthetic cycles, including sterol biosynthesis. Once the genome is completed, additional information will be forthcoming. We may find that the organism has modified its metabolic cycles to take advantage of the presence of certain host components found in the lung environment. Certainly, the repetitive genes at their telomeric ends must play a significant role in its life cycle or survival. The presence of a single rDNA locus and the organisms' inability to grow outside the mammalian lung suggest that the life style may be more host dependent than opportunistic. Perhaps we have caught *P. carinii* in the process of becoming an obligate parasite.

In contrast to the shedding of genetic material, the free-living life style of *S. cerevisiae* illustrates gene duplication and genomic polyploidization as a means for survival. Yeast must survive in harsh environments, resist plant and other toxins, and be able to respond to abundant and minimal nutrients. Thus, *S. cerevisiae* contains gene duplications for replication and growth and has evolved reproductive strategies to match the nutrient abundance.

## ACKNOWLEDGMENTS

The Pneumocystis Genome Project is supported by a grant from the National Institutes of Health, RO1AI044651.

## LITERATURE CITED

Barnett, J. A., Payne, R. W. & Yarrow, D. 1990. Yeasts: Characteristics and Identification. 2nd ed., Cambridge University Press, New York. 597 p.

Barton, E. G. J. & Campbell, W. G. J. 1967. Further observations on the ultrastructure of pneumocystis. *Arch. Pathol.*, **83**:527–534.

Barton, E. G. J. & Campbell, W. G. J. 1969. *Pneumocystis carinii* in lungs of rats treated with cortisone acetate. Ultrastructural observations relating to the life cycle. *Am. J. Pathol.*, **54**:209–236.

Campbell, W. G. J. 1972. Ultrastructure of *Pneumocystis* in human lung. Life cycle in human pneumocystosis. *Arch. Pathol.*, **93**:312–324.

Cavalier-Smith, T. 1983. A six-kingdom classification and a unified phylogeny. *In*: Schwenk, H. E. A. & Schwemmler, W.S. (ed.), Endocytobiology II: Intracellular Space as Oligogenetic. deGruyter, Berlin. p. 1027–1034.

Cavalier-Smith T. 1993. Kingdom protozoa and its 18 phyla. *Microbiol Rev.*, **57**:953–994.

Centers for Disease Control and Prevention, Division of Parasitic Diseases. Laboratory Identification of Parasites of Public Health Concern. 2003. http://www.dpd.cdc.gov/dpdx/Default.htm.

Chagas, C. 1909. Nova tripanozomiaze humana. *Mem. Inst. Oswaldo Cruz*, **1**:159–218.

Cote, R. J., Rosenblum, M., Telzak, E. E., May, M., Unger, P. D. & Cartun, R. W. 1990. Disseminated *Pneumocystis carinii* infection causing extrapulmonary organ failure: clinical, pathologic, and immunohistochemical analysis. *Mod. Pathol.*, **3**:25–30.

Cushion, M. T. 1989. In vitro studies of *Pneumocystis carinii*. *J. Protozool.*, **36**:45–52.

Cushion, M. T. 1998a. Taxonomy, genetic organization, and life cycle of *Pneumocystis carinii*. *Semin. Respir. Infect.*, **13**:304–312.

Cushion, M. T. 1998b. *Pneumocystis carinii*. *In*: Collier, L., Balows, A. & Sussman, M. (ed.), Topley and Wilson's Microbiology and Microbial Infections. Arnold & Oxford Press, New York. p. 645–683.

Cushion, M. T. 2003. *Pneumocystis*. *In*: Murray, P. R., Baron, E. J., Jorgensen, J. H., Pfaller, M. A. & Yolken, R. H. (ed.), Manual of Clinical Microbiology. ASM Press, Washington, D.C. p. 1712–1725.

Cushion, M. T. & Arnold, J. 1997. Proposal for a *Pneumocystis* genome project. *J. Eukaryot. Microbiol.*, **44**:7S.

Cushion, M. T. & Smulian, A. G. 2001. The *Pneumocystis* genome project: update and issues. *J. Eukaryot. Microbiol.*, **48**(Suppl.):182S–183S.

Cushion, M. T., Slaven, B. E. & Smulian, A. G. 2004. *Pneumocystis* genome project and genomic organization. *In*: Walzer, P. D. & Cushion, M. T. (ed.), *Pneumocystis* Pneumonia. Marcel-Dekker, Inc., New York,

Cushion, M. T., Kaselis, M., Stringer, S. L. & Stringer, J. R. 1993. Genetic stability and diversity of *Pneumocystis carinii* infecting rat colonies. *Infect. Immun.*, **61**:4801–4813

Edman, J. C., Kovacs, J. A., Masur, H., Santi, D.V., Elwood, H. J. & Sogin, M. L. 1988. Ribosomal RNA sequence shows *Pneumocystis carinii* to be a member of the Fungi. *Nature*, **334**:519–522.

Erikkson, O. E. 1994. *Pneumocystis carinii*, a parasite in the lungs of mammals referred to a new family and order (Pneumocystidaceae, Pneumocystidales, Ascomycota). *Syst. Ascomycet.*, **13**:165–180.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, C. J. & Dougherty, B. A., Merrick, J. M. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, **269**:496–512.

Frenkel, J. K. 1999. *Pneumocystis* pneumonia, an immunodeficiency-dependent disease (IDD): a critical historical overview. *J. Eukaryot. Microbiol.*, **46**:89S–92S.

Frixione, E., Ruiz, L., Santillan, M., deVargas, L. V., Tejero, J. M. & Undeen, A. H. 1992. Dynamics of polar filament discharge and sporoplasm expulsion by microsporidian spores. *Cell Motil. Cytoskelet.*, **22**:38–50.

Gimeno, C. J., Ljungdahl, P. O., Styles, C. A. & Fink, G. R. 1992. Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell*, **68**:1077–1090.

Giuntoli, D., Stringer, S. L. & Stringer, J. R. 1994. Extraordinarily low number of ribosomal RNA genes in *P. carinii*. *J. Eukaryot. Microbiol.*, **41**:88S.

Goffeau, A. et al. 1997. The Yeast Genome Directory. *Nature*, **387**(Suppl.):5–105.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S. G. 1996. Life with 6000 genes. *Science*, **274**:563–567.

Herrero, E., de la Torre, M. A. & Valentin, E. 2003. Comparative gen-

omics of yeast species: new insights into their biology. *Int. Microbiol.*, **6**:183–190.

Herskowitz, I. 1989. A regulatory hierarchy for cell specialization in yeast. *Nature,* **342**:749–757.

Hirt, R. P., Healy, B., Vossbrinck, C. R., Canning, E. U. & Embley, T. M. 1997. A mitochondrial Hsp70 orthologue in *Vairimorpha necatrix*: molecular evidence that microsporidia once contained mitochondria. *Curr. Biol.,* **7**:995–998.

Hirt, R. P., Logsdon, J. M., Healy, B., Dorey, M. W., Doolittle, W. F. & Embley, T. M. 1999. Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. *Proc. Natl. Acad. Sci. USA,* **96**:580–585.

Hughes, A. L., Friedman, R. 2003. Parallel evolution by gene duplication in the genomes of two unicellular fungi. *Genome Res.,* **13**: 794–799.

Hughes, W. T. 1987. Historical sketch. *In*: Hughes, W. T. (ed.), *Pneumocystis carinii* Pneumonitis. CRC Press, Boca Raton. p. 1–7.

Hughes, W. T., Bartley, D. L. & Smith, B. M. 1983. A natural source of infection due to *Pneumocystis carinii*. *J. Infect. Dis.,* **147**:595.

Icenhour, C. R., Rebholz, S. L., Collins, M. S. & Cushion, M. T. 2001. Widespread occurrence of *Pneumocystis carinii* in commercial rat colonies detected using targeted PCR and oral swabs. *J. Clin. Microbiol.,* **39**:3437–3441.

Icenhour, C. R., Rebholz, S. L., Collins, M. S. & Cushion, M. T. 2002. Early acquisition of *Pneumocystis carinii* in neonatal rats as evidenced by PCR and oral swabs. *Eukaryot. Cell,* 1:414–419.

Integrated Genomics I. Genomics OnLine Database (GOLD). 2003. http://ergo.integratedgenomics.com/GOLD/.

Katinka, M. D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P. & Wincker, P. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature,* **414**:450–453.

Keeling, P. J. & Fast, N. M. 2002. Microsporidia: biology and evolution of highly reduced intracellular parasites. *Annu. Rev. Microbiol.,* **56**: 93–116.

Keeling, P. J. & McFadden, G. I. 1998. Origins of microsporidia. *Trends Microbiol.,* **6**:19–23.

Keely, S. P., Wakefield, A. E., Cushion, M. T., Smulian, A. G., Hall, N., Barrell, B. G. & Stringer, J. R. 2001. Detailed structure of *Pneumocystis carinii* chromosome ends. *J. Eukaryot. Microbiol.,* **48**(Suppl.):118S–120S.

Kudo, R. R. 1947. Protozoology. Thomas. Springfield, Illinois. 778 p.

Martinsrieder Institut für Protein Sequenzen. 2003. Munich Information Centre for Protein Sequences (MIPS). Ref Type: Electronic Citation

Matsumoto, Y. & Yoshida, Y. 1984. Sporogony in *Pneumocystis carinii*: synaptonemal complexes and meiotic nuclear divisions observed in precysts. *J. Protozool.,* **31**:420–428.

Mewes, H. W., Albermann, K., Bahr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K., Kleine, K., Maierl, A., Oliver, S. G., Pfeiffer, F. & Zollner, A. 1997. Overview of the yeast genome. *Nature,* **387**: 7–8.

Prokopowich, C. D., Gregory, R. T. & Crease, T. J. 2002. The correlation between rDNA copy number and genome size in eukaryotes. *Genome,* **46**:48–50.

Slaven, B. E., Smulian, A. G. & Cushion, M. T. 2003. Pneumocystis Genome Project. http://pneumocystis.uc.edu.

Smulian, A. G., Ryan, M., Staben, C. & Cushion, M. 1996. Signal transduction in *Pneumocystis carinii*: characterization of the genes (*pcg1*) encoding the alpha subunit of the G protein (PCG1) of *Pneumocystis carinii carinii* and *Pneumocystis carinii ratti*. *Infect. Immun.* **64**:691–701.

Smulian, A. G., Sesterhenn, T., Tanaka, R. & Cushion, M. T. 2001. The *ste3* pheromone receptor gene of *Pneumocystis carinii* is surrounded by a cluster of signal transduction genes. *Genetics,* **157**:991–1002.

Stringer, J. R. & Cushion, M. T. 1998. The genome of *Pneumocystis carinii*. *FEMS Immunol. Med. Microbiol.,* **22**:15–26.

Stringer, J. R. & Keely, S. P. 2001. Genetics of surface antigen expression in *Pneumocystis carinii*. *Infect. Immun.,* **69**:627–639.

Stringer, S. L., Stringer, J. R., Blasé, M. A., Walzer, P. D. & Cushion, M. T. 1989a. *Pneumocystis carinii*: sequence from ribosomal RNA implies a close relationship with fungi. *Exp. Parasitol.,* **68**:450–461.

Stringer, S. L., Hudson, K., Blasé, M. A., Walzer, P. D., Cushion, M. T. & Stringer, J. R. 1989b. Sequence from ribosomal RNA of *Pneumocystis carinii* compared to those of four fungi suggests an ascomycetous affinity. *J. Protozool.,* **36**:14S–16S.

Taylor, J. W., Swann, E. & Berbee, M. L. 1994. Ascomycete Systematics: Problems and Perspectives in the Nineties. 1st ed., Plenum Press, New York. 421 p.

Telzak, E. E., Cote, R. J., Gold, J. W., Campbell, S. W. & Armstrong, D. 1990. Extrapulmonary *Pneumocystis carinii* infections. *Rev. Infect. Dis.,* **12**:380–386.

Thomas, C. F. J., Leof, E. B. & Limper, A. H. 1999. Analysis of *Pneumocystis carinii* introns. *Infect. Immun.,* **67**:6157–6160.

Vargas, S. L., Ponce, C. A., Gigliotti, F., Ulloa, A. V., Prieto, S., Munoz, M. P. & Hughes, W. T. 2000. Transmission of *Pneumocystis carinii* DNA from a patient with *P. carinii* pneumonia to immunocompetent contact health care workers. *J. Clin. Microbiol.,* **38**:1536–1538

Vargas, S. L., Hughes, W. T., Santolaya, M. E., Ulloa, A. V., Ponce, C. A., Cabrera, C. E., Cumsille, F. & Gigliotti, F. 2001. Search for primary infection by *Pneumocystis carinii* in a cohort of normal, healthy infants. *Clin. Infect. Dis.,* **32**:855–861.

Vavra, J. & Kucera, K. 1970. *Pneumocystis carinii delanoe*, its ultrastructure and ultrastructural affinities. *J. Protozool.,* **17**:463–483.

Vivares, C. P. & Metenier, G. 2001. The microsporidian *Encephalitozoon. Bioessays,* **23**:194–202.

Vivares, C. P., Gouy, M., Thomarat, F. & Metenier, G. 2002a. Functional and evolutionary analysis of a eukaryotic parasite genome. *Curr. Opin. Microbiol.,* **5**:499–505.

Vivares, C. P., Gouy, M., Thomarat, F. & Metenier, G. 2002b. Functional and evolutionary analysis of a eukaryotic parasitic genome. *Curr. Opin. Microbiol.,* **5**:499–505.

Vohra, P. K., Puri, V., Kottom, T. J., Limper, A. H. & Thomas, C. F. 2003. *Pneumocystis carinii* STE11, an HMG-box protein, is phosphorylated by the mitogen-activated protein kinase PCM. *Gene,* **312**: 173–179.

Vossbrinck, C. R. & Woese, C. R. 1986. Eukaryotic ribosomes that lack a 5.8S RNA. *Nature,* **320**:287–288.

Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. & Woese, C. R. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature,* **326**:411–414.

Walzer, P. D., Schnelle, V., Armstrong, D. & Rosen, P. P. 1977. Nude mouse: a new experimental model for *Pneumocystis carinii* infection. *Science,* **197**:177–179.

Weiser, J. 1985. Phylum Microspora Sprague, 1969. *In*: Lee, J. J., Hutner, S. H. & Bovee, E. C. (ed.), Illustrated Guide to the Protozoa. lst ed., Society of Protozoologists. Lawrence, Kansas. p. 375–383.

Whitehead Institute MIoT. 2003. The Fungal Genome Initiative. http://www-genome.wi.mit.edu/annotation/fungi/fgi/.

Wolfe, K. H. & Shields, D. C. 1997. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature,* **387**:708–713.

Wright, J. H. & Craighead, E. M. 1922. Infectious motor paralysis in young rabbits. *J. Exp. Med.,* **36**:135–149.

Ye, D., Lee, C. H. & Queener, S. F. 2001. Differential splicing of *Pneumocystis carinii* f. sp. *carinii* inosine 5′-monophosphate dehydrogenase pre-mRNA. *Gene,* **263**:151–158.

*Received 08/22/2003; accepted 08/23/2003*