MINIREVIEWS

Simplicity and Complexity of Microsporidian Genomes

Patrick J. Keeling* and Claudio H. Slamovits

*Botany Department, Canadian Institute for Advanced Research, University of British
Columbia, Vancouver, British Columbia, Canada*

The adaptation to a parasitic lifestyle typically precipitates a number of profound changes to various levels of biological organization, including cell structure, genetics and genomics, metabolism, and biochemistry. These changes can lead to a seemingly paradoxical mixture of characteristics: on one hand, parasites may evolve extremely complex and sophisticated mechanisms to invade their host and evade its defenses, while on the other hand, they may also appear more "simple" by dispensing with characteristics they no longer need as they increasingly depend on host metabolism for nutrients and energy. The reductive element of parasitic evolution has led several parasitic protozoa to be considered ancient or early-branching eukaryotes because they lack characters common to all other eukaryotes, but in reality, these organisms are highly derived. No group represents this phenomenon better than the microsporidia, a group of obligate intracellular parasites that exemplify the contradictory forces of simplification and complex specialization.

MICROSPORIDIA—SIMPLE YET COMPLEX

Microsporidia are a large and diverse group of intracellular parasites (66). Over 1,000 species of microsporidia have been described to date (36), but this is likely a small fraction of the true diversity of the group, which might be closer to the number of species of animals. Virtually all known microsporidia infect animals; they are known to infect members of every animal phylum. Some species infect multiple distantly related hosts, but specificity to one host or a related group of hosts seems to be the norm for the group (50, 55). Over a dozen microsporidia, most of which are opportunistic, are known to infect humans (16, 35).

The complexity of microsporidian cells rests in their unique and highly adapted infection machinery (60). Microsporidia are obligate intracellular parasites, so they must infect other eukaryotic cells to grow and divide. The agents of infection are highly resistant spores, which are the only stages of the life cycle that can survive in the environment. Spores contain a number of complex structures that mediate infection: a posterior vacuole, an anterior lamella of membranes called the polaroplast, and a long, coiled filament called the polar filament

(33, 60). When induced to germinate, the spore builds up osmotic pressure until its rigid wall ruptures at its thinnest point at the apex. The posterior vacuole swells, forcing the polar filament to rapidly eject from the rupture, simultaneously everting itself to form a tube. The contents of the spore are forced through the tube, and if it has penetrated another cell during its rapid ejection, the parasite is injected into the cytoplasm of this potential host (33, 59).

Surely such a cell cannot be considered simple, but in many other ways, it is quite so. Indeed, aside from organelles of infection, the spore contains a nucleus or two, cytoplasm, and little else. Spores and intracellular stages lack, or were thought to lack, mitochondria, stacked Golgi dictyosomes, flagella and other 9 + 2 microtubular structures, and peroxisomes. Moreover, they have a highly reduced metabolism, 70S ribosomes, and among the smallest known genomes of any eukaryote (4, 28, 60, 62). The simple side of microsporidia led to the idea that they may be very primitive eukaryotes, and they were once considered part of a group of protists called archezoa, which were believed to include eukaryotes that evolved prior to the acquisition of mitochondria (9). In addition to microsporidia, archezoa included diplomonads (e.g., *Giardia* spp.), parabasalia (e.g., *Trichomonas* spp.), and archamoebae (e.g., *Entamoeba* spp.). Initially, molecular data supported the idea that these organisms belonged to ancient, amitochondriate lineages (27, 63), but more recently, this conclusion has been overturned. In the case of the microsporidia, two crucial facts that show that they are highly derived, rather than primitively simple and amitochondriate, have emerged. First, molecular phylogenetics have shown that microsporidia did not evolve early in eukaryotic evolution (i.e., they are not primitive), despite what the first trees showed. Instead, it is now clear that they are either fungi or close relatives of the fungi (12, 15, 24, 28, 29, 31, 32). Second, microsporidia do contain mitochondria or a relict version of the organelle called a mitosome. The first indications of this fact came from the discovery of genes for mitochondrion-derived proteins in the nuclear genomes of several microsporidia (13, 18, 23, 44), and subsequently, the complete genome of *Encephalitozoon cuniculi* revealed over a dozen mitochondrion-derived protein-coding genes (28). Eventually, the organelle itself was identified by immunolocalization of mitochondrial HSP70 in *Trachipleistophora hominis* (64), eliminating the last shred of doubt that the microsporidia evolved from a mitochondriate ancestor and retained a relict, highly reduced organelle.

Today the Archezoa hypothesis has been undermined for each of the four original members of the group, and our per-

* Corresponding author. Mailing address: Canadian Institute for Advanced Research, Botany Department, University of British Columbia, 3529-6270 University Blvd., Vancouver, BC V6T 1Z4, Canada. Phone: (604) 822-4906. Fax: (604) 822-6089. E-mail: pkeeling @interchange.ubc.ca.
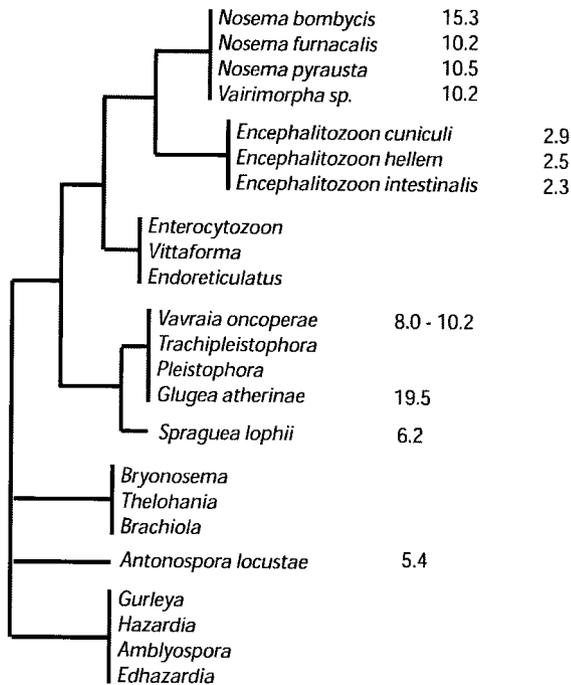
FIG. 1. Schematic tree depicting evolutionary relationships among representative microsporidian lineages. Resolved branching points represent nodes that are strongly supported in SSU rRNA trees (8, 17, 45, 53). The position of *Nosema pyrausta* is from the work of Baker et al. (1). Numbers are genome size estimates in millions of base pairs (see reference 38 for a review).

ception of microsporidia has been turned completely upside down: while they were for a time regarded as primitively simple eukaryotes, we now see them as highly adapted, specialized parasites derived from complex fungal ancestors. This new perspective colors much of our interpretation of microsporidian biology: where we once saw primitive characteristics, we now see reduction. Nowhere is this transition in our thinking more relevant than in the genomics of microsporidia.

REDUCTION BY LOSS AND COMPACTION IN MICROSPORIDIAN GENOMES

The process of reduction is most tangible in the genomes of microsporidia. Karyotypic analyses have shown microsporidian genomes to be typical in their overall structure (i.e., multiple linear chromosomes) but very small (Fig. 1). They range in size from a relatively unremarkable 19.5 Mbp in *Glugea atherinae* (5) to only 2.3 Mbp, one of the smallest eukaryotic genomes known, in *Encephalitozoon intestinalis* (43). For perspective, the *E. intestinalis* genome is over 1,000 times smaller than the human genome (3,000 Mbp), only one-quarter the size of the *Saccharomyces cerevisiae* genome (12 Mbp), which is itself considered small for a eukaryote, and almost half the size of the *Escherichia coli* genome (4.6 Mbp). So what makes a microsporidian genome small? There are two main ways in which a genome can be reduced: it can reduce the number of genes it encodes, or it can pack them into a smaller space. These two processes are likely driven by different forces, and they have

different implications for the genome form, function, and content (30).

The first complete microsporidian genome, that of *E. cuniculi*, has been a turning point for microsporidian molecular biology, providing valuable information about the biochemistry and molecular physiology of the parasite, as well as the nature of the genome itself (28). The *E. cuniculi* genome consists of 11 linear chromosomes ranging from 217 to 315 kb, totaling only 2.9 Mb. Chromosomes are made up of gene-rich cores flanked by subtelomeric repeats consisting of rRNA operons and telomeres. The use of rRNA operons in subtelomeric repeats is interesting, as it has been found in a number of other eukaryotic genomes, including many that are compact. This arrangement is found in the ultracompact nucleomorph genomes of endosymbionts of the chlorarachniophyte *Bigelowiella natans* and the cryptophyte *Guillardia theta* (11, 19, 68), and there is evidence for the arrangement in the microsporidian *Antonospora locustae* (53).

The coding potential of the *E. cuniculi* genome is very low, with fewer than 2,000 protein-encoding genes being recognized (28), showing that substantial reduction by gene loss has likely taken place. The distribution of the *E. cuniculi* genes among functional categories shows that gene loss is not random. Instead, absent genes represent complete metabolic or regulatory pathways (28, 61). As the organism becomes host dependent, some metabolic functions are provided by the host cell, so genes involved in such functions may become redundant and lost after inactivation by mutation. This process is illustrated by the distribution of protein-encoding genes in *E. cuniculi*: proteins involved in biosynthesis of small molecules are highly underrepresented (28, 61), while genes predicted to encode numerous transporters that import energy sources (e.g., glucose and ATP), other metabolites (e.g., nucleosides and amino acids), and ions are found. Not surprisingly, pathways related to many basic cellular processes, such as DNA replication, transcription, and protein synthesis, are conserved more or less intact (38).

These characteristics show that the *E. cuniculi* genome has been reduced by a loss of genes, but what about compaction? The density of the gene-rich cores is approximately one gene for every 1,000 bp (28), which is the highest gene density among the completely sequenced eukaryotic nuclear genomes. For comparison, *S. cerevisiae* is generally considered to have a very compact genome at one gene for every 2,000 bp (21). It is noteworthy that the density of the *E. cuniculi* genome is similar to that of the nucleomorph genome of *G. theta* (11) and may represent a maximum gene density for nuclear genomes.

The packing of genes into a smaller space has taken place in two distinct ways in *E. cuniculi*. First, many of the genes themselves are small. One way that this quality is achieved is that all but a few introns have been eliminated from the genome (only 13 remain). More surprisingly, however, about 80% of the predicted proteins themselves are 15% smaller on average than their homologues in yeast (28). Katinka and coworkers (28) proposed that the loss of proteins and pathways resulted in fewer functional interactions between proteins (i.e., simpler protein interaction networks), which in turn allowed the loss of domains involved in these interactions from many other proteins that were retained. This shortening of proteins is an interesting and unanticipated effect of genome reduction that
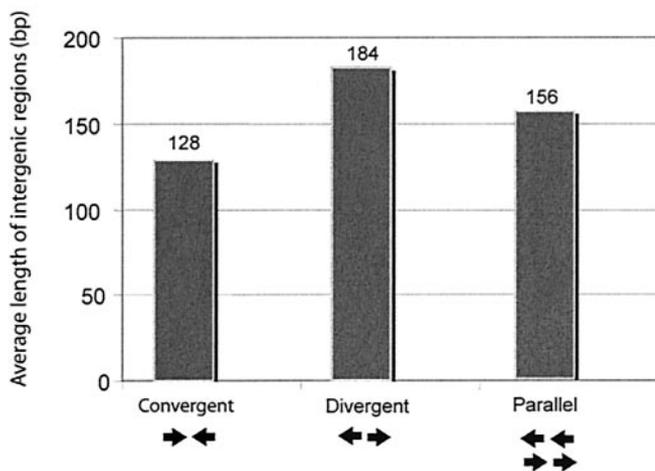
FIG. 2. Genome compaction and intergenic DNA. Bars represent average intergenic regions in base pairs of *E. cuniculi* for the three possible orientations of two genes. The graph is based on 1,771 intergenic regions, with data from the work of Katinka et al. (28). Repeated genes and short, subtelomeric ORFs were excluded.

has a number of functional implications which need further study, but it accounts for only a fraction of genome compaction: if *E. cuniculi* proteins were the same sizes as their yeast homologues, the *E. cuniculi* genome would be only about 0.25 Mbp bigger. Most of the reduction is achieved by the loss of intergenic DNA. Not surprisingly, *E. cuniculi* is devoid of highly repeated (dispersed and satellite) DNA (28), which accounts for a large fraction of the genome in many eukaryotes (34, 41). There are only a few duplicated segments and subtelomeric repeats of about 20 kb at the chromosome ends. Likewise, no evidence of active or relict mobile elements has been found, although there is limited evidence for them in other microsporidia (22, 39). Nonrepetitive intergenic DNA can also constitute a significant fraction of eukaryotic genomes because it includes the sometimes-complex arrays of regulatory elements essential to promote and regulate transcription. Broadly, promoters are composed of a core where the basal transcription apparatus binds with different degrees of affinity, together with additional elements that direct the binding of regulatory transcription factors. In spite of extensive work in this area, the evolution of promoters and transcriptional regulation is only partly understood (67). One feature of eukaryotic promoters that is known to be important is the spacing between regulatory binding sites (65), and overall, the amount of intergenic DNA involved in transcriptional regulation has been estimated to be at least equal to the number of functional coding sites in nematodes (48) and mammals (49). In sharp contrast, the intergenic regions of *E. cuniculi* are very short, averaging only 129 bp (28). Interestingly, the lengths of intergenic spaces vary noticeably, depending on the transcriptional orientation of the two flanking genes (Fig. 2). Genes that are transcribed divergently from one another have 50% more intergenic DNA than genes arranged convergently, while genes in the same orientation have an intermediate amount of DNA. This pattern is just detectable in other eukaryotic genomes but quite pronounced in the *E. cuniculi* genome, likely because much of the "noise" has been removed from noncoding re-

gions. Since most regulatory elements are found 5′ from a gene, the observed pattern is expected when the amount of intergenic sequence is under severe reductive selection.

COMPARATIVE GENOMICS AND THE IMPACTS OF GENOME REDUCTION IN MICROSPORIDIA

Clearly, extreme reduction of gene numbers has an impact on the metabolic and functional versatility of the cell, but what effects does genome compaction have on the function of the genome itself? One way to look at this problem is to use comparative genomics to examine both the shared characteristics among different microsporidia and the dynamics of the genome over time. Three genome sequence surveys have been carried out with microsporidian species other than *E. cuniculi*. The genomes of the opportunistic human pathogen *Vittaforma corneae* and the fish parasite *Spraguea lophii* have been subject to small surveys (22, 39). The most significant outcome was the finding of a reverse transcriptase gene in both species, suggesting the presence of retrotransposable elements. However, these studies looked only at relatively short genomic fragments and did not provide clues to the influence of such elements in genome architecture. *A. locustae* (formerly *Nosema locustae*) (53, 54) is another microsporidian for which more-substantial genomic data have now been produced (13, 14, 51). *A. locustae* has been shown to be relatively distantly related to *E. cuniculi* (53, 54); however, a random genome survey of 685 kbp showed that the two genomes have a good deal in common. First, of all the genes identified in *A. locustae* to date, a large proportion (130 out of 138, or 94%) are also found in the *E. cuniculi* genome. Three genes have homologues in other organisms but are absent in *E. cuniculi*, and five known *A. locustae* open reading frames (ORFs) have little or no similarity to anything else and could represent either highly divergent genes or *A. locustae*-specific genes. Considering the apparent evolutionary distance separating these species and the differences in their life cycles and hosts, such conservation in gene content is rather surprising.

The similarity between these genomes extends beyond gene content, however, to the actual order of the genes (51). Gene order conservation is typically lost relatively quickly in eukaryotic genomes, mostly by frequent short inversions (26, 47). However, in many cases, genes were found in the same order in the *A. locustae* and *E. cuniculi* genomes, and in one case this conservation can also be inferred from a third microsporidian, *Spraguea lophii* (Fig. 3). Indeed, the overall degree of synteny between *A. locustae* and *E. cuniculi*, two distantly related microsporidia, is about 1.5 times higher than that between *S. cerevisiae* and *Candida albicans*, two closely related ascomycetes, and is also higher than that between humans and fish (51). In contrast, the rate of small-subunit (SSU) rRNA divergence between *A. locustae* and *E. cuniculi* is 10 times higher than that of *S. cerevisiae* and *C. albicans*, and protein divergence is almost twice as high (Fig. 4). *Saccharomyces* and *Candida* are thought to have diverged about 200 million years ago (2), or about the time placental mammals diverged from marsupials. It is very unlikely that microsporidia are younger than these yeasts, because they are known to infect all phyla of animals in all ecological settings. For microsporidia to have evolved so recently, they would have had to infect some animals and then
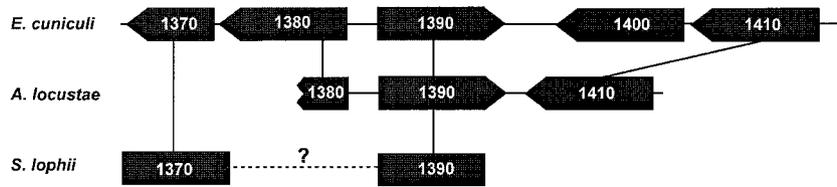
FIG. 3. Example of gene order conservation among three microsporidian species. An *E. cuniculi* fragment of five genes shows partial conservation with an *A. locustae* fragment and two *Spraguea lophii* genes located in the same genomic clone (only the ends of the clone are known, so the dashed line indicates a region of unknown length and sequence). Labels on the genes correspond to the *E. cuniculi* locus denominations (28). The internal sequence of the *Spraguea lophii* clone and the relative orientations of both genes are unknown. *Spraguea lophii* data were taken from http://jbpc.mbl.edu/Spraguea-HTML/.

quickly spread through the entire kingdom, which contrasts with the typically restricted host range of individual microsporidia. The alternative explanation is simply that microsporidian genomes are rearranging slowly, in contrast with their rapid sequence evolution. Enhanced genome stability has been observed in several other systems and explained in a number of ways. Cotranscription of conserved gene clusters and lack of recombination are two of the more common ways (25, 46, 57, 58). Interestingly, it has been proposed that the reduction of intergenic distances will enhance genome stability simply by making nondeleterious breakpoints rare, reducing the likelihood of inversions or transpositions (25). This effect is detectable in yeasts but has been shown to be a relatively insignificant component compared with cotranscription (25). In the extremely compacted genomes of microsporidia, however, it is possible that what is a minor force in "normal" genomes takes on a more substantial role, resulting in the greater degree of stability apparent between these two distantly related species. The absences of repetitive and mobile DNA in the genome further enhance this effect, although evidence for some repetitive DNA exists in *A. locustae* (14), and retroelements exist in *V. corneae* (39) and *Spraguea lophii* (22).

If compaction has enhanced genome stability in microsporidia, then it has probably done so only at certain levels of organization. This theory is likely true because chromosome size polymorphisms have been documented in two well-studied microsporidian genomes, *E. cuniculi* and *Encephalitozoon hellem* (3, 6, 7). Indeed, those studies led to the idea that the *Encephalitozoon* genome is rather plastic (38), which contrasts with the evidence for increased genome stability at the gene
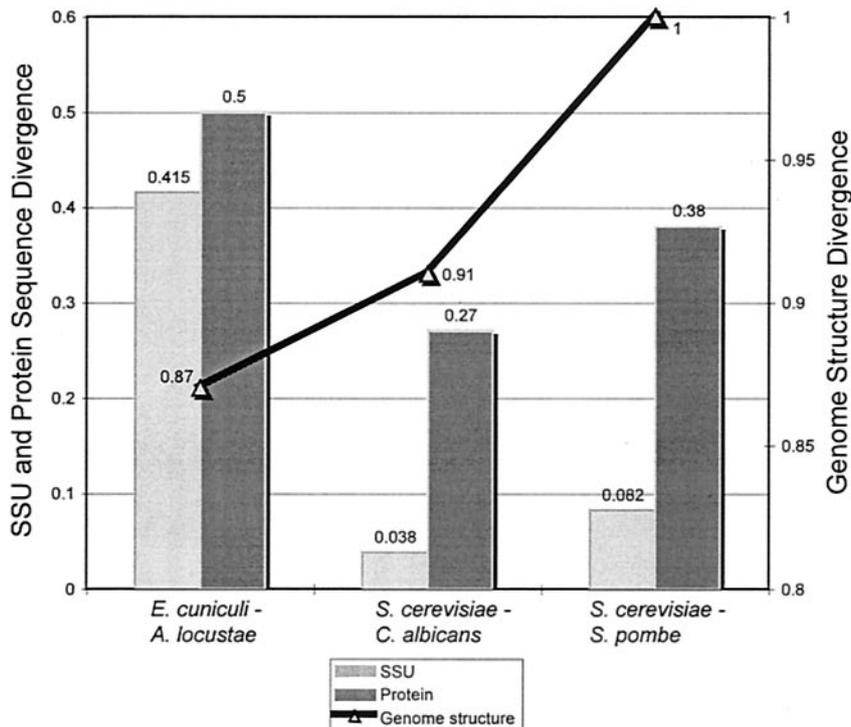

FIG. 4. Comparing genome structures and degrees of sequence evolution in microsporidia and fungi. Bars represent degrees of sequence divergence for SSU rRNA nucleotides (light grey) and a variety of proteins at the amino acid level (dark grey) for the pairs *E. cuniculi-A. locustae*, *S. cerevisiae-C. albicans*, and *S. cerevisiae-Schizosaccharomyces pombe*. The black line represents the degree of genome evolution (rearrangement). Values are 1 minus the proportions of shared adjacent genes for each pair of species, which are 0.13, 0.09, and 0, respectively (data are from reference 51).

order level. If the conservation of gene order is a result of compaction, it is likely that the gene-rich chromosome cores will be stable, whereas the gene-poor areas may be subject to a high plasticity. Consistent with this idea, rearrangements involved in chromosomal size polymorphisms in *E. cuniculi* seem to take place in subtelomeric regions adjacent to rRNA gene units where duplications take place (7). However, these differences in size are too small (typically 2 to 12 kbp) to significantly influence the overall levels of genome compaction seen in *E. cuniculi*, and whether such polymorphism exists in other species is not known.

WHAT IS THE DIFFERENCE BETWEEN MICROSPORIDIAN GENOMES?

The *E. cuniculi* genome is among the smallest of all microsporidia, whereas the largest is nearly seven times its size (Fig. 1). While there is significant diversity in microsporidian life cycles and environments, there is no obvious difference in complexity to account for such a drastic variation. This observation raises the obvious question, what makes the difference between differently sized microsporidian genomes? This question has been posed many times for genomes in general (10, 37, 42), but the microsporidia present a unique subset of the question because of their apparent selection for compaction and reduction. Is the variation in genome size among microsporidia a product of different gene densities, different numbers of genes, or both, and if they are all compacted, how can this be reconciled with huge differences in genome size?

Currently, there is very little genomic data from microsporidia with variously sized genomes. The *A. locustae* genome provides the best comparison so far, as it is estimated to be nearly twice the size of the *E. cuniculi* genome (5.4 Mbp) (56), and yet it has a very similar complement of genes (51). How, then, can we explain the difference in genome size? One possibility is that the estimated genome size of *A. locustae* is wrong, but pulsed-field electrophoresis generally provides a reasonable minimum estimate for genome size, and the *E. cuniculi* genome has proven to be very close to its estimated size (4). Assuming that the estimated size is correct, there are two main and nonexclusive alternatives: the genome is less compact and/or it contains more genes.

If the *A. locustae* genome (and the larger genomes of other microsporidia) is less compact, it must contain proportionately more noncoding DNA than *E. cuniculi*. Sequence data from the *A. locustae* genomic survey suggest that this idea may be true, but only to a point. Most characterized *A. locustae* genome fragments have gene densities very close to that of *E. cuniculi* (0.94 in *A. locustae* versus 0.97 in *E. cuniculi*). A few regions, however, are less gene dense. One atypical 13-kbp fragment was found to encode only three ORFs separated by comparatively long intergenic distances (3.8 and 1.9 kb). A 500-bp segment of one of these intergenic regions contains five 84-bp repeats, and there is evidence that these repeats are present in multiple copies in the genome (14). These observations suggest that some variation exists in the densities of microsporidian genomes and that repeated noncoding DNA might be more common in *A. locustae* than it is in *E. cuniculi*, but this suggestion probably represents an insignificant difference overall.

So what about the gene complement of *A. locustae*? Because the genome is not complete, it is impossible to state that *E. cuniculi* has genes that *A. locustae* lacks, but it is possible to identify *A. locustae* genes that are absent from *E. cuniculi*. The current data on genome size and density suggest that it should have hundreds of genes that are absent from *E. cuniculi*, but as mentioned earlier, this also does not seem to be the case. Of the 138 *A. locustae* genes currently characterized, only 8 do not have clear homologues in *E. cuniculi*, and most of these are unidentified ORFs. A few genes with obvious identities have been found in *A. locustae* but not in *E. cuniculi*, and some of these are very interesting. One of the first such genes to be found was a class II catalase that has been shown to be active in the spore and was proposed to protect spores against oxidative stress (14). Interestingly, phylogenetic analyses showed that this gene was acquired relatively recently from a proteobacterium by lateral gene transfer. A similar situation is seen in the *A. locustae* gene for photolyase, an enzyme which mediates the repair of UV-damaged DNA, using visible light as an energy source. Photolyase is absent from *E. cuniculi*, but the *A. locustae* gene is expressed in spores and may complement photolyase-deficient *E. coli* (52). The distribution of photolyase in microsporidia is completely unknown but may be interesting since the *Antonospora* protein is of a class common to animals but so far not found in fungi. It is possible that it also arose by lateral gene transfer, but the phylogeny does not resolve this issue. Regardless of how photolyase originated, the presence of catalase and photolyase in *A. locustae*, but not in *E. cuniculi*, does show that the former has an enhanced repertoire of stress response proteins compared with those of *E. cuniculi*, leading to the prediction that *E. cuniculi* spores are less resistant to various kinds of environmental damage (14, 52). Moreover, these examples show the potential for variation among contents of microsporidian genomes with functionally interesting implications, and since only about 10% of the *A. locustae* genome has been examined in detail, there are many more such examples to be discovered. As interesting as each of these may be, however, the overall number of genes present in *A. locustae* but not *E. cuniculi* is still not sufficient to explain the differences in their respective genome sizes, even if it is (unrealistically) assumed that *A. locustae* does not lack any genes which *E. cuniculi* has.

FUTURE PERSPECTIVES—A GENOMIC BIODIVERSITY VIEW OF MICROSPORIDIA

The *E. cuniculi* genome has set the stage for comparative studies that will redefine not only how we see microsporidian genomes but also how we see processes and forces that act on highly specialized and reduced eukaryotic genomes in general. How, for instance, can a genome appear to have the same gene complement and the same gene density but be almost twice as large? More dramatically, what is the difference between the largest microsporidian genomes and the smallest? It seems doubtful that they have seven times as many unique genes, and given the state of the *A. locustae* genome, it would also be surprising if their genomes were significantly less compact. Do the genomes reveal a fundamental disconnect between the forces of compaction and size reduction, two processes we instinctively assumed to be one and the same? Unfortunately,

there are no obvious answers to these riddles at present, but the existing data hint at the potential for a revolution in the way that we look at reduced genomes.

With respect to the microsporidia specifically, the current data also teach us a lesson often forgotten in the age of model organisms: a greater sampling of genome content is bound to break down many of our assumptions about what a group of organisms is like. The small sampling of genomic data from *A. locustae* has already shown that, despite the many similarities between it and *E. cuniculi*, there is also much variability, and this variability will become tremendously more apparent if even a few species are examined in more detail. In addition, the apparently high degree of synteny will be much more interesting when more dimensions are added to the data. For example, are the same gene clusters preserved across most microsporidian genomes, or do we simply see about the same degree of conservation overall? Such differences may seem subtle, but they reveal important aspects of the processes behind these observations: in this example, the distinction would reveal whether conserved gene clusters were somehow special as opposed to part of a more random but global process.

Lastly, reductive forces acting on eukaryotic genomes may be best understood when microsporidian genomes are systematically compared with those of other eukaryotes where the process has been taken to similar or even greater extremes. The most notable of these are the small genomes of picoeukaryotes (40) and the nucleomorph genomes of cryptomonads (11) and chlorarachniophytes (19, 20). The latter are particularly interesting as they have been reduced nearly to extinction but still retain a few hundred kilobases. As all these systems have evolved independently and likely for very different reasons, it would be rash to assume that they all reflect the same forces. Nevertheless, there are bound to be some common elements to the way these genomes have responded to apparently harsh selective pressures, pressures we are only beginning to understand.

REFERENCES

1. **Baker, M. D., C. R. Vossbrinck, E. S. Didier, J. V. Maddox, and J. A. Shadduck.** 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. J. Eukaryot. Microbiol. **42:**564–570.

2. **Berbee, M. L., and J. W. Taylor.** 2001. Fungal molecular evolution: gene trees and geologic time. *In* D. J. McLaughlin, E. G. McLaughlin, and P. A. Lemke (ed.), The mycota, vol. VII, part B. Springer-Verlag, Berlin, Germany.

3. **Biderre, C., A. Mathis, P. Deplazes, R. Weber, G. Méténier, and C. P. Vivarès.** 1999. Molecular karyotype diversity in the microsporidian *Encephalitozoon cuniculi.* Parasitology **118:**439–445.

4. **Biderre, C., M. Pages, G. Méténier, E. U. Canning, and C. P. Vivarès.** 1995. Evidence for the smallest nuclear genome (2.9 Mb) in the microsporidium *Encephalitozoon cuniculi.* Mol. Biochem. Parasitol. **74:**229–231.

5. **Biderre, C., M. Pages, G. Méténier, D. David, J. Bata, G. Prensier, and C. P. Vivarès.** 1994. On small genomes in eukaryotic organisms: molecular karyotypes of two microsporidian species (Protozoa) parasites of vertebrates. C. R. Acad. Sci. Ser. III **317:**399–404.

6. **Brugere, J. F., E. Cornillot, T. Bourbon, G. Méténier, and C. P. Vivarès.** 2001. Inter-strain variability of insertion/deletion events in the *Encephalitozoon cuniculi* genome: a comparative KARD-PFGE analysis. J. Eukaryot. Microbiol. **48**(Suppl.):50S–55S.

7. **Brugere, J. F., E. Cornillot, G. Méténier, and C. P. Vivarès.** 2000. Occurrence of subtelomeric rearrangements in the genome of the microsporidian parasite *Encephalitozoon cuniculi,* as revealed by a new fingerprinting procedure based on two-dimensional pulsed field gel electrophoresis. Electrophoresis **21:**2576–2581.

8. **Canning, E. U., D. Refardt, C. R. Vossbrinck, B. Okamura, and A. Curry.** 2002. New diplokaryotic microsporidia (phylum Microsporidia) from freshwater bryozoans (Bryozoa, Phylactolaemata). Eur. J. Protistol. **38:**247–265.

9. **Cavalier-Smith, T.** 1983. A 6-kingdom classification and a unified phylogeny, p. 1027–1034. *In* H. E. A. Schenk and W. S. Schwemmler (ed.), Endocytobiology, vol. II. Walter de Gruyter & Co., Berlin, Germany.

10. **Cavalier-Smith, T.** 1985. The evolution of genome size. John Wiley and Sons, Chichester, United Kingdom.

11. **Douglas, S., S. Zauner, M. Fraunholz, M. Beaton, S. Penny, L. T. Deng, X. Wu, M. Reith, T. Cavalier-Smith, and U. G. Maier.** 2001. The highly reduced genome of an enslaved algal nucleus. Nature **410:**1091–1096.

12. **Edlind, T. D., J. Li, G. S. Visversvara, M. H. Vodkin, G. L. McLaughlin, and S. K. Katiyar.** 1996. Phylogenetic analysis of the β-tubulin sequences from amitochondriate protozoa. Mol. Phylogenet. Evol. **5:**359–367.

13. **Fast, N. M., and P. J. Keeling.** 2001. Alpha and beta subunits of pyruvate dehydrogenase E1 from the microsporidian *Nosema locustae:* mitochondrion-derived carbon metabolism in microsporidia. Mol. Biochem. Parasitol. **117:**201–209.

14. **Fast, N. M., J. S. Law, B. A. P. Williams, and P. J. Keeling.** 2003. Bacterial catalase in the microsporidian *Nosema locustae:* implications for microsporidian metabolism and genome evolution. Eukaryot. Cell **2:**1069–1075.

15. **Fast, N. M., J. M. Logsdon, Jr., and W. F. Doolittle.** 1999. Phylogenetic analysis of the TATA box binding protein (TBP) gene from *Nosema locustae:* evidence for a microsporidia-fungi relationship and spliceosomal intron loss. Mol. Biol. Evol. **16:**1415–1419.

16. **Franzen, C., and A. Muller.** 2001. Microsporidiosis: human diseases and diagnosis. Microbes Infect. **3:**389–400.

17. **Fries, I., R. J. Paxton, J. Tengö, S. B. Slemenda, A. J. da Silva, and N. J. Pieniazek.** 1999. Morphological and molecular characterization of *Antonospora scoticae* n. gen., n. sp. (Protozoa, Microsporidia) a parasite of the communal bee, *Andrena scotica* Perkins, 1916 (Hymenoptera, Andrenidae). Eur. J. Protistol. **35:**183–193.

18. **Germot, A., H. Philippe, and H. Le Guyader.** 1997. Evidence for loss of mitochondria in Microsporidia from a mitochondrial-type HSP70 in *Nosema locustae.* Mol. Biochem. Parasitol. **87:**159–168.

19. **Gilson, P., and G. I. McFadden.** 1995. The chlorarachniophyte: a cell with two different nuclei and two different telomeres. Chromosoma **103:**635–641.

20. **Gilson, P. R., and G. I. McFadden.** 2002. Jam packed genomes—a preliminary, comparative analysis of nucleomorphs. Genetica **115:**13–28.

21. **Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver.** 1996. Life with 6000 genes. Science **274:**546, 563–567.

22. **Hinkle, G., H. G. Morrison, and M. L. Sogin.** 1997. Genes coding for reverse transcriptase, DNA-directed RNA polymerase, and chitin synthase from the microsporidian *Spraguea lophii.* Biol. Bull. **193:**250–251.

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

24. **Hirt, R. P., J. M. Logsdon, Jr., B. Healy, M. W. Dorey, W. F. Doolittle, and T. M. Embley.** 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.

25. **Hurst, L. D., E. J. Williams, and C. Pal.** 2002. Natural selection promotes the conservation of linkage of co-expressed genes. Trends Genet. **18:**604–606.

26. **Huynen, M. A., B. Snel, and P. Bork.** 2001. Inversions and the dynamics of eukaryotic gene order. Trends Genet. **17:**304–306.

27. **Kamaishi, T., T. Hashimoto, Y. Nakamura, F. Nakamura, S. Murata, N. Okada, K.-I. Okamoto, M. Shimizu, and M. Hasegawa.** 1996. Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. J. Mol. Evol. **42:**257–263.

28. **Katinka, M. D., S. Duprat, E. Cornillot, G. Méténier, F. Thomarat, G. Prenier, V. Barbe, E. Peyretaillade, P. Brottier, P. Wincker, F. Delbac, H. El Alaoui, P. Peyret, W. Saurin, M. Gouy, J. Weissenbach, and C. P. Vivarès.** 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi.* Nature **414:**450–453.

29. **Keeling, P. J.** 2003. Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. Fungal Genet. Biol. **38:**298–309.

30. **Keeling, P. J.** 2004. Reduction and compaction in the genome of the apicomplexan parasite, *Cryptosporidium parvum.* Dev. Cell **6:**614–616.

31. **Keeling, P. J., and W. F. Doolittle.** 1996. Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Mol. Biol. Evol. **13:**1297–1305.

32. **Keeling, P. J., M. A. Luker, and J. D. Palmer.** 2000. Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. Mol. Biol. Evol. **17:**23–31.

33. **Keohane, E. M., and L. M. Weiss.** 1999. The structure, function, and composition of the microsporidian polar tube, p. 196–224. *In* M. Wittner and L. M. Weiss (ed.), The microsporidia and microsporidiosis. ASM Press, Washington, D.C.

34. **Kidwell, M. G.** 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica **115:**49–63.

35. **Kotler, D. P., and J. M. Orenstein.** 1999. Clinical syndromes associated with

microsporidiosis. *In* M. Wittner and L. M. Weiss (ed.), The microsporidia and microsporidiosis. ASM Press, Washington, D.C.

36. **Larsson, J. I. R.** 1999. Identification of microsporidia. Acta Protozool. **38:** 161–197.

37. **Lynch, M., and J. S. Conery.** 2003. The origins of genome complexity. Science **302:**1401–1404.

38. **Méténier, G., and C. P. Vivarès.** 2001. Molecular characteristics and physiology of microsporidia. Microbes Infect. **3:**407–415.

39. **Mittleider, D., L. C. Green, V. H. Mann, S. F. Michael, E. S. Didier, and P. J. Brindley.** 2002. Sequence survey of the genome of the opportunistic microsporidian pathogen, *Vittaforma corneae*. J. Eukaryot. Microbiol. **49:**393–401.

40. **Moreira, D., and P. Lopez-Garcia.** 2002. The molecular ecology of microbial eukaryotes unveils a hidden world. Trends Microbiol. **10:**31–38.

41. **Ohno, S.** 1972. So much "junk" DNA in our genome. *In* H. H. Smith (ed.), Evolution of genetic systems. Gordon and Breach, New York, N.Y.

42. **Petrov, D. A.** 2001. Evolution of genome size: new approaches to an old problem. Trends Genet. **17:**23–28.

43. **Peyretaillade, E., C. Biderre, P. Peyret, F. Duffieux, G. Méténier, M. Gouy, B. Michot, and C. P. Vivarès.** 1998. Microsporidian *Encephalitozoon cuniculi*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. Nucleic Acids Res. **26:**3513–3520.

44. **Peyretaillade, E., V. Broussolle, P. Peyret, G. Méténier, M. Gouy, and C. P. Vivarès.** 1998. Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. Mol. Biol. Evol. **15:**683–689.

45. **Refardt, D., E. U. Canning, A. Mathis, S. A. Cheney, N. J. Lafranchi-Tristem, and D. Ebert.** 2002. Small subunit ribosomal DNA phylogeny of microsporidia that infect Daphnia (Crustacea: Cladocera). Parasitology **124:** 381–389.

46. **Saccone, C., C. Gissi, A. Reyes, A. Larizza, E. Sbisa, and G. Pesole.** 2002. Mitochondrial DNA in metazoa: degree of freedom in a frozen event. Gene **286:**3–12.

47. **Seoighe, C., N. Federspiel, T. Jones, N. Hansen, V. Bivolarovic, R. Surzycki, R. Tamse, C. Komp, L. Huizar, R. W. Davis, S. Scherer, E. Tait, D. J. Shaw, D. Harris, L. Murphy, K. Oliver, K. Taylor, M. A. Rajandream, B. G. Barrell, and K. H. Wolfe.** 2000. Prevalence of small inversions in yeast gene order evolution. Proc. Natl. Acad. Sci. USA **97:**14433–14437.

48. **Shabalina, S. A., and A. S. Kondrashov.** 1999. Pattern of selective constraint in *C. elegans* and *C. briggsae* genomes. Genet. Res. **74:**23–30.

49. **Shabalina, S. A., A. Y. Ogurtsov, V. A. Kondrashov, and A. S. Kondrashov.** 2001. Selective constraint in intergenic regions of human and mouse genomes. Trends Genet. **17:**373–376.

50. **Shaw, R. W., M. L. Kent, A. M. Brown, C. M. Whipps, and M. L. Adamson.** 2000. Experimental and natural host specificity of *Loma salmonae* (Microsporidia). Dis. Aquat. Org. **40:**131–136.

51. **Slamovits, C. H., N. M. Fast, J. S. Law, and P. J. Keeling.** 2004. Genome compaction and stability in microsporidian intracellular parasites. Curr. Biol. **14:**891–896.

52. **Slamovits, C. H., and P. J. Keeling.** 2004. Class II photolyase in a microsporidian intracellular parasite. J. Mol. Biol. **341:**713–721.

53. **Slamovits, C. H., B. A. Williams, and P. J. Keeling.** 2004. Transfer of *Nosema locustae* (Microsporidia) to *Antonospora locustae* n. comb. based on molecular and ultrastructural data. J. Eukaryot. Microbiol. **51:**207–213.

54. **Sokolova, Y. Y., V. V. Dolgikh, E. V. Morzhina, E. S. Nassonova, I. V. Issi, R. S. Terry, J. E. Ironside, J. E. Smith, and C. R. Vossbrinck.** 2003. Establishment of the new genus *Paranosema* based on the ultrastructure and molecular phylogeny of the type species *Paranosema grylli* Gen. Nov., Comb. Nov. (Sokolova, Selezniov, Dolgikh, Issi 1994), from the cricket *Gryllus bimaculatus* Deg. J. Invertebr. Pathol. **84:**159–172.

55. **Solter, L. F., and J. V. Maddox.** 1998. Physiological host specificity of microsporidia as an indicator of ecological host specificity. J. Invertebr. Pathol. **71:**207–216.

56. **Streett, D. A.** 1994. Analysis of *Nosema locustae* (Microsporidia: Nosematidae) chromosomal DNA with pulsed-field gel electrophoresis. J. Invertebr. Pathol. **63:**301–303.

57. **Suyama, M., and P. Bork.** 2001. Evolution of prokaryotic gene order: genome rearrangements in closely related species. Trends Genet. **17:**10–13.

58. **Tamas, I., L. Klasson, B. Canback, A. K. Naslund, A. S. Eriksson, J. J. Wernegreen, J. P. Sandstrom, N. A. Moran, and S. G. Andersson.** 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science **296:**2376–2379.

59. **Undeen, A. H., and R. K. Vander Meer.** 1999. Microsporidian intrasporal sugars and their role in germination. J. Invertebr. Pathol. **73:**294–302.

60. **Vávra, J., and J. I. R. Larsson.** 1999. Structure of the microsporidia, p. 7–84. *In* M. Wittner and L. M. Weiss (ed.), The microsporidia and microsporidiosis. ASM Press, Washington, D.C.

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

62. **Vivarès, C. P., and G. Méténier.** 2000. Towards the minimal eukaryotic parasitic genome. Curr. Opin. Microbiol. **3:**463–467.

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

64. **Williams, B. A., R. P. Hirt, J. M. Lucocq, and T. M. Embley.** 2002. A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. Nature **418:**865–869.

65. **Wingender, E.** 1993. Gene regulation in eukaryotes. VCH, Weinheim, Germany.

66. **Wittner, M., and L. M. Weiss.** 1999. The microsporidia and microsporidiosis. ASM Press, Washington, D.C.

67. **Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman, and L. A. Romano.** 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. **20:**1377–1419.

68. **Zauner, S., M. Fraunholz, J. Wastl, S. Penny, M. Beaton, T. Cavalier-Smith, U. G. Maier, and S. Douglas.** 2000. Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph. Proc. Natl. Acad. Sci. USA **97:**200–205.