

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/49739238>

Shrink it or lose it: Balancing loss of function with shrinking genomes in the microsporidia

Article in *Virulence* · January 2011

DOI: 10.4161/viru.2.1.14606 · Source: PubMed

CITATIONS

19

READS

79

2 authors, including:



Nicolas Corradi

University of Ottawa

184 PUBLICATIONS **2,151** CITATIONS

SEE PROFILE

Shrink it or lose it

Balancing loss of function with shrinking genomes in the microsporidia

Patrick J. Keeling^{1*} and Nicolas Corradi²

¹Canadian Institute for Advanced Research; Botany Department; University of British Columbia; Vancouver, BC; ²Department of Biology; University of Ottawa; Ottawa, ON Canada

Microsporidia are obligate intracellular parasites that have evolved an elaborate mechanism for invading animal host cells, but which have otherwise greatly reduced biological complexity. In particular, microsporidia possess the smallest autonomous nuclear genomes known (as opposed to nucleus derived organelles or nucleomorphs), and their 'anaerobic' core carbon metabolism is severely limited. Here we compare the extremes to which these two characteristics have evolved, and contrast how their reduction has either proceeded within the constraints of an unchanging set of functions, or has reduced the functional capabilities of the cell. Specifically, we review how the smallest known nuclear genome, the 2.3 Mbp genome of *Encephalitozoon intestinalis*, has arrived at this diminutive form without significantly affecting its protein-coding complexity in comparison with closely related, larger genomes. In contrast to this, *Enterocytozoon bieneusi* has a relatively large genome, and yet has lost all enzymes necessary to synthesize ATP from sugar—imposing a major limitation on the functional capabilities of the cell. The extremity of this reduction demands a re-evaluation of metabolic processes in other microsporidia: although pathways such as glycolysis are present, comparative genomic data suggest they may not play the cellular role that they are generally assumed to play.

Introduction

Microsporidia are an abundant and diverse lineage of obligate intracellular

parasites related to fungi.^{1,2} Outside their host cell these organisms are only viable as highly resistant spores that are equipped with an intriguing suite of sub-cellular structures used to gain entry to their host cell. Most obvious of these is the polar filament, a coil wound about the inside of the spore that is ejected at high speed during germination, inverting at the same time to become a tube. Should this projectile happen to penetrate the cytoplasm of a host, it allows the contents of the spore to move through the tube and into that cytoplasm, where it can begin its infectious cycle.³ This mechanism may be used to attack a host cell from outside, to move between adjacent host cells or to escape from an endocytic vacuole to the cytoplasm.⁴

Aside from the machinery that mediates infection, microsporidia are surprisingly simple: so simple, in fact, they were for a time argued to be ancient, primitive eukaryotes.⁵ Based on molecular phylogenies, we now know this is not the case; instead, they are related to fungi and as such we now know they are derived or reduced, from more complex ancestors.² This reduction has been the focus of some attention due to the extremes to which microsporidia have gone in losing characteristics we normally expect in a eukaryote, and it is now clear that these strange parasites have taken several routes to this end. Here, two different ways to reduce biological complexity that have been recently described in three papers⁶⁻⁸ will be illustrated and, by contrasting these, some of their broader implications will be discussed.

Key words: microsporidia, parasite, genomics, metabolism, evolution, subtelomere, fungi, anaerobic, glycolysis

Submitted: 10/28/10

Revised: 12/09/10

Accepted: 12/10/10

DOI: 10.4161/viru.2.1.14606

*Correspondence to: Patrick J. Keeling;
Email: pkeeling@interchange.ubc.ca

Addendum to: Corradi N, Pombert JF, Farinelli L, Didier ES, Keeling PJ. The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. *Nat Commun* 2010; 1:77; PMID: 20865802; DOI: 10.1038/ncomms1082.

and Keeling PJ, Corradi N, Morrison HG, Haag KL, Ebert D, Weiss LM, et al. The reduced genome of the parasitic microsporidian *Enterocytozoon bieneusi* lacks genes for core carbon metabolism. *Genome Biol Evol* 2010; 2:304-9; PMID: 20624735; DOI: 10.1093/gbe/evq022.

Loss of DNA—Trimming the Fat from Skinny Genomes

Though there are many odd things about the biology of microsporidia, the small size of their genomes has stood out since their genomes were first investigated.^{9,10} It is hard to make concrete generalizations about nuclear genomes, but we do have an imperfect concept of how they differ from those of bacteria: they are composed of multiple linear chromosomes with centromere and telomeres, and they are more often than not large by comparison. In structure, microsporidian genomes fit this conception, but in content many of them do not. The largest microsporidian genomes are 25–30 Mbp,^{10,11} well within the range for microbial eukaryotes, but others are much smaller, and indeed are the smallest known nuclear genomes of any eukaryote. At the extreme is the genus *Encephalitozoon*, a widespread parasite of vertebrates, including humans, where genome sizes smaller than 3 Mbp have been measured.^{10,12} To put this into context, this is more than 1,000x smaller than the human genome, and almost twice as small as that of *E. coli* K12. Almost ten years ago the genome of *Encephalitozoon cuniculi* was completely sequenced, and found to be a mere 2.9 Mbp.¹³ The genome has become a model for reduced and compacted genomes, with about 2,000 genes and a bacterial-like gene density.

But *E. cuniculi* does not have the smallest microsporidian genome—the smallest currently known is found in its cousin, *E. intestinalis*, which was estimated to be 2.3 Mbp or 20% smaller.¹² Looking at the low number of genes and extreme density of the *E. cuniculi* genome, one wonders, where is the slack? If *E. intestinalis* really is 20% smaller, what is it missing compared to *E. cuniculi*?

To find out, we sequenced the complete genome of *E. intestinalis*.⁷ The sequence, composed of 11 chromosomes with no internal gaps, conformed closely to the original estimate of 2.3 Mbp and showed quite clearly where the 20% difference between *E. cuniculi* and *E. intestinalis* lay. Between the chromosome ‘cores’ (the region defined as spanning between the first and last identifiably homologous genes in the two species) there was little

difference—more or less the same genes, in the same order, and at a similar density. Although the chromosome cores represent the vast majority of the genome, they only accounted for only 34 kbp of the size difference, suggesting it has been reduced to the limit of its functional constraints. This conclusion was supported by sequence evolution as well: intergenic regions showed a lower rate of substitution than synonymous positions in the genes surrounding them—suggesting most of the remaining intergenic sequences were constrained by selection, and so were likely functionally significant, or in other words if there was any excess, it had already been removed.⁷

The chromosome ends, on the other hand, could hardly have been more different. The sequenced sub-telomeric regions of the *E. intestinalis* chromosomes were all significantly smaller than the corresponding ends in *E. cuniculi*, so that a total of at least 224 kbp encoding 174 genes were absent from *E. intestinalis*.⁷

Broadly, this answers the question of how an already compressed genome can be further reduced—simply snip off the ends of all the chromosomes. Sounds easy, but it raises as many questions as it answers. *E. cuniculi* is no slouch in terms of genome reduction, so why does it have hundreds of genes at its chromosome ends if they are dispensable in a closely related species? Clearly they are, under similar circumstances, non-essential, so why would such a tightly-packed genome have many kilobases of such genes, and why are they all clustered at the chromosome ends? The nature of the genes in question is obviously central to this, and not surprisingly many of them are unidentified ORFs or part of repeated families or both. It could be that these genes and families are rapidly turned over in these genomes, and *E. cuniculi* just happens to represent a temporary ‘bloated’ state due to a recent expansion.⁷ To determine if the genes were gained in *E. cuniculi* or lost in *E. intestinalis*, the genome of a third species will be necessary. It is also worth noting that that this difference may be adaptation in response to some pressure (i.e., there is some advantage to having a small genome if you are an *Encephalitozoon*). While intuitive, this has not been demonstrated convincingly, and it is also possible that

even extreme reduction is due to some intrinsic characteristic of the genome that is non-adaptive. For instance, a bias towards deletions over insertions with limited recombination could pack the genome very tightly, and in small populations (the population structure of various microsporidia remains an open question) this could allow even mildly deleterious changes to drift to fixation.

As interesting as these changes are for genome evolution at its most radical extremes, it is noteworthy that they potentially have very little effect on the molecular or cell biology of the organisms themselves because of the nature of the missing genes and the fact that many of those that are part of duplicated families with homology elsewhere in the *E. intestinalis* genome. Overall, *E. intestinalis* is not missing any known pathway or functional cluster of genes that is present in *E. cuniculi*,⁷ so this difference in genome size could be the result of an intrinsic property of the genome itself as opposed to a reaction to any external pressure or change.

Loss of Function—The Deceptive Simplicity of Microsporidian Metabolism

In contrast to the trimming of the *E. intestinalis* genome, other microsporidia have taken other reductive routes just as far; and here metabolism stands out. Microsporidian metabolism is at the best of times very restricted. Direct biochemical analyses are hampered by their obligate intracellular nature,¹⁴ but in the species where we have abundant sequence data, the metabolism can be modeled based on what genes are present.^{15,16} In all such species, many small molecule biosynthetic pathways are completely absent, and core carbon metabolism is limited to glycolysis, the pentose phosphate pathway, and synthesis and degradation of the storage carbohydrate trehalose—even the smallest *E. intestinalis* genome has the full complement of genes for these pathways.⁷ In *Enterocytozoon bieneusi*, however, virtually all these genes have been lost. This species cannot be maintained indefinitely *in vitro*,¹⁷ so sequencing its entire genome is a challenge, and without a complete genome it is difficult to conclude a gene has really

been lost. To get around this problem, two separate, deep genome sequence surveys have been conducted on *E. bienersi* and the list of genes generated by these two surveys compared.^{8,18} Remarkably, nearly identical lists of genes were found, suggesting the genome is very well sampled, though incomplete. More specifically, an analysis of over 300 genes representing 30 functional pathways suggested that over 90% of the genes found in *E. cuniculi* had also been sampled in *E. bienersi*. In contrast, of the 21 *E. cuniculi* genes for core carbon metabolism, only two were found in *E. bienersi* and, significantly, the same two were found in both surveys.^{8,18} Moreover, fatty acid metabolism was also severely underrepresented and, fascinatingly, both spliceosomal introns and all the proteins and RNAs associated with spliceosomal splicing were completely absent, suggesting this genome has also lost all traces of intron splicing.^{8,18} Overall, it appears that this species completely lacks any capacity to carry out several normal eukaryotic activities, the most surprising of which is the ability to generate energy from sugars. Instead, it must rely on importing ATP directly from its host. Unlike *E. intestinalis*, the reductive process in *E. bienersi* has not led to a particularly small genome (it is thought to be about 6 Mbp¹⁸) but, instead, to a major change of function since what are normally important functional pathways have been entirely lost.

The absence of these pathways in *E. bienersi* has striking implications for its metabolism, but also significant implications on the role played by these pathways in other microsporidia. Glycolysis is generally regarded as the universal backbone of energy metabolism in anaerobic eukaryotes; and microsporidia are no different.^{19,20} However, although the pathway is intact in other microsporidian genomes, the *E. bienersi* genome and other findings described below give us reason to doubt its role in these species. In addition to generating ATP, glycolysis generates reducing potential in the form of NADH. Without an outlet for this potential, it must be dissipated or the pathway becomes energetically unfavorable. Some eukaryotes dissipate this excess potential by transferring electrons to the mitochondrion via the glycerol-3-phosphate (G3PDH)

shuttle where they are absorbed by the alternative oxidase (AOX).^{21,22} Recently, homologs of AOX have been found in some, but interestingly not all, microsporidia.⁶ *Antonosporea locustae*, for example, has genes for all the enzymes of glycolysis, AOX, both partners for the G3PDH shuttle, and proteins necessary to target the mitochondrial G3PDH to the mitochondrion;^{6,23,24} altogether making up a complete system to make ATP from glucose (aerobically, which is noteworthy). Not surprisingly, Enterocytozoon has none of these systems intact,⁸ but other species such as Encephalitozoon, fall somewhere between these two extremes. Encephalitozoon genomes lack AOX⁶ and the G3PDH shuttle has been shown to be broken in *E. cuniculi*, since the mitochondrial partner is localized to the cytosol²⁵ and it lacks a peptidase responsible for processing the targeted product in other eukaryotes.²⁴ Overall, Encephalitozoon might have glycolysis but apparently lacks any obvious capacity to keep the pathway going continuously, as one might imagine it would do if it were being used as a significant source of ATP.

This suggests Encephalitozoon metabolism may be closer to that of Enterocytozoon than their genomes superficially suggest. Encephalitozoon genomes, and indeed all known microsporidian genomes, encode several ATP transporters,²⁶ some that import ATP from the host,²⁷ and others that import it into their reduced mitochondrion.²⁸ So even if a given species of microsporidian is not entirely dependent on host ATP, as Enterocytozoon seems to be, all microsporidia appear to at least have the capacity to be so dependent. Perhaps the real questions should not be why *E. bienersi* has lost glycolysis, but rather why all the other microsporidia that possess ATP transporters have not lost it as well? If ATP can be imported, then the synthesis of ATP (via glycolysis or any other means) needs to be explained in the context of some particular life stage or environmental clue making it critical for survival.

Lacking the necessary components to keep glycolysis running continuously, it appears possible that Encephalitozoon does not actually use glycolysis for energy: like Enterocytozoon, Encephalitozoon

may also import all its ATP from its host. If this is true, and Encephalitozoon lacks the ability to use glycolysis as a continuous source of energy, then why has it retained all the genes for this pathway when they have been lost in Enterocytozoon? One possibility is, ironically, that Encephalitozoon is using glycolysis to make NADH. While excess NADH is a potential problem with glycolysis running continuously, a small amount of NADH could be made by metabolizing glucose (sparingly) by glycolysis, and the ATP also produced would be a useful by-product of this NADH production.

Concluding Remarks— Microsporidia Entering the Era of Comparative Genomics

Microsporidia are a phylogenetically diverse group of parasites and, although they share the same fundamental infection mechanism, they also display a significant amount of diversity in how they use that mechanism, how they affect their hosts, their life cycle variations, and in other traits relating to their life histories.²⁹ Indeed, their genetic diversity and ubiquity across the diversity of animals leads one to speculate that they may have originated at about the same time as the animal lineage and represent a nearly equivalent amount of diversity. Although impossible to quantify, anecdotal data suggest that any animal species investigated in sufficient detail will eventually be found to be infected by one of more species of microsporidia and that many of these have limited host range. One averages, one could even mount a reasonable argument that there are a similar number of species of microsporidian as there are of animals. For these reasons, we have paid a great deal of attention to the variation between different microsporidia, but it is worth pointing out that one of the emerging themes of comparative genomics of microsporidia is the remarkable lack of variation at other levels. Despite the differences detailed above, the overall gene content of diverse microsporidia is remarkably homogeneous.^{7,8,11,13,18,30,31} Indeed, the overall degree of conservation in gene content serves to highlight what differences we do find. Nevertheless, these differences

are real and many are of functional or evolutionary significance, so comparative genomics has also served to remind us that no one 'model' will adequately represent the lineage as a whole.

Comparing the genomes of microsporidia has also served to illustrate that the genome of one species may be difficult to interpret at face value. The presence of genes for a complete glycolytic pathway in *E. cucurbitae* can present an obvious interpretation on its own, but superimposing on this the implications of findings from the *E. bienersi* and *A. locustae* genomes significantly alters this interpretation. Given the relatively small size of even the largest microsporidian genomes, next-generation sequencing technology will soon enable the generation of a great deal of comparative data from many microsporidian species: this kind of data has already changed much of how we understand this group (e.g., many of them should no longer be considered 'anaerobes' as they once were described because they metabolize oxygen using AOX) and how this transition continues will be interesting to observe. One level where such technology will be of obvious interest is that of the population. The population structure of most microsporidia is all but unknown, so even basic questions such as whether many species are sexual or not are mysteries. Just as comparing genomes between distantly related microsporidian genera or between different species of the same genus addressed distinct kinds of questions, so too comparing genomes from multiple strains of a microsporidian species has the potential to tell us a great deal about other levels of diversity, evolution and even protein and genome function.

Acknowledgements

Our work on microsporidian genomics is supported by a grant from the Canadian Institutes for Health Research (MOP-42517). P.J.K. is a Senior Scholar of the Michael Smith Foundation for Health

Research and a Fellow of the Canadian Institute for Advanced Research. N.C. is a Scholar of the Canadian Institute for Advanced Research.

References

- Keeling P. Five questions about microsporidia. *PLoS Pathog* 2009; 5:1000489.
- Corradi N, Keeling PJ. Microsporidia: a journey through radical taxonomic revisions. *Fung Biol Rev* 2009; 23:1-8.
- Keoghane EM, Weiss LM. The structure, function and composition of the microsporidian polar tube. In: Wittner M, Weiss LM, Eds. *The microsporidia and microsporidiosis*. Washington: ASM Press 1999; 196-224.
- Franzen C. How do microsporidia invade cells? *Folia Parasitol (Praha)* 2005; 52:36-40.
- Cavalier-Smith T. A 6-kingdom classification and a unified phylogeny. In: Schenk HEA, Schwemmler WS, Eds. *Endocytobiology II: Intracellular Space as Oligogenetic*. Berlin New York: Walter de Gruyter and Co 1983; 1027-34.
- Williams BA, Elliot C, Burri L, Kido Y, Kita K, Moore AL, et al. A broad distribution of the alternative oxidase in microsporidian parasites. *PLoS Pathog* 2010; 6:1000761.
- Corradi N, Pombert JF, Farinelli L, Didier ES, Keeling PJ. The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. *Nat Commun* 2010; 1:77.
- Keeling PJ, Corradi N, Morrison HG, Haag KL, Ebert D, Weiss LM, et al. The reduced genome of the parasitic microsporidian *Encephalitozoon bienersi* lacks genes for core carbon metabolism. *Genome Biol Evol* 2010; 2:304-9.
- Bidre C, Pagès M, Méténier G, Canning EU, Vivarès CP. Evidence for the smallest nuclear genome (2.9 Mb) in the microsporidium *Encephalitozoon cucurbitae*. *Mol Biochem Parasitol* 1995; 74:229-31.
- Bidre C, Pagès M, Méténier G, David D, Bata J, Prensier G, et al. On small genomes in eukaryotic organisms: molecular karyotypes of two microsporidian species (Protozoa) parasites of vertebrates. *C R Acad Sci III* 1994; 317:399-404.
- Corradi N, Haag KL, Pombert JF, Ebert D, Keeling PJ. Draft genome sequence of the Daphnia pathogen *Octospora bayeri*: insights into the gene content of a large microsporidian genome and a model for host-parasite interactions. *Genome Biol* 2009; 10:106.
- Peyretaille E, Bidre C, Peyret P, Duffieux F, Méténier G, Gouy M, et al. Microsporidian *Encephalitozoon cucurbitae*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. *Nucleic Acids Res* 1998; 26:3513-20.
- Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prensier G, et al. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cucurbitae*. *Nature* 2001; 414:450-3.
- Weidner E, Findley AM, Dolgikh V, Sokolova J. Microsporidian biochemistry and physiology. In: Wittner M, Weiss LM, Eds. *The Microsporidia and Microsporidiosis*. Washington DC: ASM Press 1999.
- Vivarès CP, Gouy M, Thomarat F, Méténier G. Functional and evolutionary analysis of a eukaryotic parasitic genome. *Curr Opin Microbiol* 2002; 5:499-505.
- Méténier G, Vivarès CP. Molecular characteristics and physiology of microsporidia. *Microbes Infect* 2001; 3:407-15.
- Vivaresva GS. In vitro cultivation of microsporidia of clinical importance. *Clin Microbiol Rev* 2002; 15:401-13.
- Akiyoshi DE, Morrison HG, Lei S, Feng X, Zhang Q, Corradi N, et al. Genomic survey of the non-cultivable opportunistic human pathogen, *Encephalitozoon bienersi*. *PLoS Pathog* 2009; 5:1000261.
- van der Giezen M, Tovar J. Degenerate mitochondria. *EMBO Rep* 2005; 6:525-30.
- Williams BAP, Keeling PJ. Cryptic organelles in parasitic protists and fungi. *Adv Parasitol* 2003; 54:9-67.
- Chaudhuri M, Ott RD, Hill GC. Trypanosome alternative oxidase: from molecule to function. *Trends Parasitol* 2006; 22:484-91.
- Suzuki T, Hashimoto T, Yabu Y, Kido Y, Sakamoto K, Nihei C, et al. Direct evidence for cyanide-insensitive quinol oxidase (alternative oxidase) in apicomplexan parasite *Cryptosporidium parvum*: phylogenetic and therapeutic implications. *Biochem Biophys Res Commun* 2004; 313:1044-52.
- Williams BA, Keeling PJ. Microsporidian mitochondrial proteins: expression in *Antonospora locustae* spores and identification of genes coding for two further proteins. *J Eukaryot Microbiol* 2005; 52:271-6.
- Burri L, Williams BAP, Bursac D, Lithgow T, Keeling PJ. Microsporidian mitochondria retain elements of the general mitochondrial targeting system. *Proc Natl Acad Sci USA* 2006; 103:15916-20.
- Williams BA, Cali A, Takvorian PM, Keeling PJ. Distinct localization patterns of two putative mitochondrial proteins in the microsporidian *Encephalitozoon cucurbitae*. *J Eukaryot Microbiol* 2008; 55:131-3.
- Richards TA, Hirt RP, Williams BA, Embley TM. Horizontal gene transfer and the evolution of parasitic protozoa. *Protist* 2003; 154:17-32.
- Tsaousis AD, Kunji ER, Goldberg AV, Lucocq JM, Hirt RP, Embley TM. A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cucurbitae*. *Nature* 2008; 453:553-6.
- Williams BA, Haferkamp I, Keeling PJ. An ADP/ATP-specific mitochondrial carrier protein in the microsporidian *Antonospora locustae*. *J Mol Biol* 2008; 375:1249-57.
- Vávra J, Larsson JIR. Structure of the Microsporidia. In: Wittner M, Weiss LM, Eds. *The Microsporidia and Microsporidiosis*. Washington DC: American Society for Microbiology 1999; 7-84.
- Cornman RS, Chen YP, Schatz MC, Street C, Zhao Y, Desany B, et al. Genomic analyses of the microsporidian *Nosema ceranae*, an emergent pathogen of honey bees. *PLoS Pathog* 2009; 5:1000466.
- Slamovits CH, Fast NM, Law JS, Keeling PJ. Genome compaction and stability in microsporidian intracellular parasites. *Curr Biol* 2004; 14:891-6.