Available online at www.sciencedirect.com
ScienceDirect
Genomics 91 (2008) 388–393
GENOMICS
www.elsevier.com/locate/ygeno

Comparative profiling of overlapping transcription in the compacted genomes of microsporidia *Antonospora locustae* and *Encephalitozoon cuniculi* ☆

Nicolas Corradi, Anna Gangaeva, Patrick J. Keeling *

Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, Canada BC V6T 1Z4

Received 25 September 2007; accepted 17 December 2007
Available online 20 February 2008

Abstract

Microsporidia are highly adapted parasites related to fungi with compact, gene-dense genomes. It has previously been shown in the microsporidian *Antonospora locustae* that transcripts from any given gene overlap with adjacent genes at a high frequency, perhaps due to the compact nature of its genome. However, it is still not known if this phenomenon is widespread among microsporidia or conserved between species, or even whether it is strictly correlated with compaction. To address these questions, we performed a comparison of transcription profiles in two microsporidian species, *A. locustae* and *Encephalitozoon cuniculi.* Transcription overlap was characterized at many *A. locustae* loci representing a range of gene densities, to determine if overlapping transcription correlates with the length of intergenic spacers. In parallel, we examined the first cases of transcription overlap in *E. cuniculi.* Using regions of the genome where the order of genes is conserved between *A. locustae* and *E. cuniculi,* we identified the transcriptional processing points in both species to determine how the process changes through evolutionary time. We show that there is little conservation of processing points between species and indeed that the process differs in important ways in the two genomes. Overall, *A. locustae* transcripts generally start just upstream of the start codon, but terminate well within or beyond downstream genes. In contrast, *E. cuniculi* transcripts often initiate within upstream genes, but more frequently terminate prior to the downstream gene. This process appears to have predictable characteristics within a given genome, but to be relatively flexible between species, presenting further challenges to the study of gene expression in these obligately intracellular parasites.

Keywords: Overlapping transcription; *Antonospora locustae*; *Encephalitozoon cuniculi*; Genome compaction; Processing points

Microsporidia are a large and diverse group of obligate intracellular eukaryotic parasites consisting of about 160 genera and 1300 currently described species [1]. They are especially common in arthropods and fish [2], but are increasingly recognized as parasites of mammals, including humans. Microsporidia share a number of primitive characteristics that led to the hypothesis that they were among the earliest lineages of eukaryotes [3]. However, it is now widely acknowledged that they are specialized and derived relatives of fungi [4–6].

The nuclear genomes of microsporidia tend to be small, and in many cases extremely small. The genome of *Encephalitozoon intestinalis* is only 2.3 Mb, about half the size of that of *Escherichia coli* K12, and many other species have genomes less than 10 Mb [7,8]. The smallest genomes are the best studied, and they are characterized by both reduction and compaction. The complete genome of *Encephalitozoon cuniculi* (2.9 Mb) contains only about 2000 genes and has a gene density of about 1 gene/kb, roughly twice the density of yeast [5,9]. It has been recently shown that such hypercompaction can have consequences on its transcriptional machinery. Specifically, a high proportion of transcripts from the microsporidian *Antonospora locustae* and the similarly compacted nucleomorph genomes from two algal endosymbionts were found to encode

☆ Sequence data from this article have been deposited with the GenBank Data Library under Accession Nos. ES880821 to ES880896.

* Corresponding author. Fax: +1 604 822 6089.
E-mail address: pkeeling@interchange.ubc.ca (P.J. Keeling).

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.12.006

fragments of more than one gene [10]. This was concluded to be due to the fact that promoters and terminators no longer fit in the most highly reduced intergenic regions. This phenomenon, known as overlapping, or multigene transcription, is not identical to operons in prokaryotes [11], since only one gene appears to be expressed (a notion supported by the observation that many gene fragments on a transcript are encoded on opposite strands) [12]. Overlapping transcription is not unknown in other genomes (e.g., transcripts from certain loci have been shown to overlap in fungi [13,14]), but what sets *A. locustae* (and nucleomorphs) apart is the high frequency with which this appears to be tolerated.

Because overlapping transcription was detected in *A. locustae* through an expressed sequence tag (EST) survey, we have a broad picture of transcription across that genome, but several questions remain outstanding. Here we examine two of these questions. First, the analyses of EST sequences revealed few cases in which multiple cDNAs from the same overlapping loci could be compared, and many cDNAs were 5′ truncated [12]. Overall, therefore, little is known about the frequency of overlap at the 5′ end and how variable initiation and polyadenylation sites are. Without such information, it is not certain if the EST data reflect a relaxed control of transcription in compacted genomes, such that transcription starts and stops at variable positions around a gene because promoters and terminators are "weak" and transcription sloppy. Alternatively, promoters and terminators may be strong and well organized, but often located at specific positions within adjacent genes. In this case, transcripts should have consistent and well-defined initiation and termination sites. These alternatives have different implications for the effects of overlapping transcription in microsporidian genomes, but current data fail to distinguish between these hypotheses. Second, the extent and variability of overlapping transcription within microsporidia are both unknown, and by extension it is unclear if its occurrence contributes to genome structure. The completely sequenced genome of *E. cuniculi* [9] has been found to share high levels of gene order conservation with *A. locustae,* despite the fact that the two species are only distantly related [15,16]. More recently analysis of parts of the *Nosema bombycis* and *Enterocytozoon bieneusi* genomes have also revealed conservation in genome order [17,18], suggesting that the genome has changed slowly throughout microsporidian evolution. This slow pace of change has been attributed to compaction, but overlapping transcription could also play an obvious role: if the control elements for a gene are embedded within an adjacent gene, separating the two genes could often be deleterious.

Here, we compare transcriptional processing in *A. locustae* and *E. cuniculi.* First, we developed a comprehensive transcriptional end-point map for a number of loci containing expressed transcripts in *A. locustae* [10] and from these derived an overall pattern of transcription in this species. Second, we identified the locations of all *E. cuniculi* homologues encoded within *A. locustae* overlapping transcripts [10] and found that the gene order is conserved in several previously unrecognized cases. Targeting these regions of the genome and other regions with short intergenic spaces, we determined the transcriptional processing points for many homologous loci from both species. Overall these data show that the nature of transcripts is different in the two species in important ways, that the length of the intergenic regions is a strong indicator of the occurrence of overlapping transcription, and that processing points vary greatly between organisms, suggesting transcriptional overlap does not greatly contribute to genome stability.

Results

Sampling of multigene transcripts in A. locustae

A total of 17 *A. locustae* loci were analyzed using 3′- and 5′-RACE to determine the overall degree of variation in transcription start and end points in this species. Loci separated by both short and large intergenic regions were selected, and also cases in which adjacent genes were encoded on the same strand or the opposite strand (Table 1, Supplemental Fig. 1). Consistent with most of the previous observations from this species [10,12], the transcription initiation sites in all but a few loci were very close to the translation initiation codon; in many cases the A of the AUG appeared to be capped. In no case was transcription initiation observed within an upstream gene in the present data, as has been reported for photolyase [12]. 3′-RACE fragments from *A. locustae* were more variable, and the frequency of overlapping fragments was roughly the same as the frequency estimated from EST data [10]. In some cases, however, multiple polyadenylation sites were identified for a gene (never corresponding to an A track in the genome, so we conclude these are not DNA-derived contaminants), suggesting that several processing signals can be present within a single downstream gene or intergenic region.

Table 1
Summary of initiation and polyadenylation positions in *A. locustae* transcripts

Gene name	Initiation site[a]	5′ overlap	Poly(A) site[b]	3′ overlap
Actin	N/D	N/D	61	None
Hypothetical protein	5	None	>312[c]	None
Polyubiquitin	N/D	N/D	N/D	N/D
VID25	0	None	345/491	Polyubiquitin
WD-repeat protein	4	None	N/D	N/D
Hypothetical protein	1	None	>122[c]	Cytidine deaminase
Cytidine deaminase	N/D	N/D	673	None
Checkpoint kinase	5	None	211	PDH-E1α
Phosphoesterase	N/D	N/D	374	None
Internalin	6	None	10/262	None
Lipoprotein	0	None	74	None
p68-like protein	1	None	104	None
Phosphate transporter	0	None	299	None
Lea	0	None	74	None
Rpl9	0	None	83	None
Sno RNA-associated protein	9	None	38	None
DNA-binding protein	1	None	258	None
CTP synthase	0	None	14	None

All data are available as Supplemental Fig. 1.
[a]Base pairs upstream of first base of ATG.
[b]Base pairs downstream of last base of TAA/TAG/TGA.
[c]Data from 5′-RACE of downstream gene.

Comparing transcription in E. cuniculi and A. locustae

To determine if overlapping transcription is broadly distributed among microsporidia, and to determine the similarities and differences in the process between distantly related species, we selected a number of *E. cuniculi* loci to examine and compare with homologous loci in *A. locustae.* Plausible candidates were selected from loci where overlapping transcription is already known in *A. locustae* or loci with short intergenic regions where the gene order is shared between *A. locustae* and *E. cuniculi.*

In four cases in which overlapping transcripts in *A. locustae* were identified by EST sequencing, we found the gene order was also conserved in *E. cuniculi* (Fig. 1, Supplemental Fig. 2). Analysis of the 5′ and 3′ ends of *E. cuniculi* transcripts of one or more genes in each of these regions demonstrated that most also overlap with adjacent loci in this species. In only two cases could the same end of the same gene be compared: at the 3′ end of the 20S proteasome subunit α gene there is overlap with *rpl6* in *A. locustae* but not *E. cuniculi,* and the 3′ end of *cdc2* overlaps with the downstream ring zinc finger gene in *A. locustae* but not in *E. cuniculi* (Fig. 1). Overall, the pattern of transcript processing observed in *E. cuniculi* is very different from that of *A. locustae.* In contrast to *A. locustae,* in which transcription almost always begins at or near the AUG codon, in *E. cuniculi* transcription most frequently initiates within the upstream gene (Fig. 1, Supplemental Fig. 2). Moreover, in some cases, 5′-RACE yielded products of different sizes, which appear to encode transcripts that start at different positions. In contrast, 3′-RACE fragments from *E. cuniculi* were mostly found to terminate a few nucleotides downstream of the stop codon, with a single exception that ended in a downstream gene (tRNA adenylyl transferase, for which the intergenic space is only 9 bp; Supplemental Fig. 2).

To compare directly the same ends of the same genes in more detail, we identified four other genomic regions where *E. cuniculi* and *A. locustae* shared the same overall gene order (Fig. 2). We analyzed 5′ and 3′ transcript ends from one or more genes from each of these regions from both genomes, resulting in 12 direct comparisons. In general, the nature of these transcripts was in keeping with the trends noted above for their respective genomes. More specifically, there were six cases in which the 3′ ends were identified from homologues from both species, and in all six they shared the same state (two with overlap, four without). Of the six 5′ ends identified from homologues from both species, only two were the same (both nonoverlapping) and the remainder differed between the two species (in all of which there was an overlap in *E. cuniculi,* but not in *A. locustae*).

Discussion

Transcription overlap is correlated with genome compaction

Among the factors that contributed to the compaction of the *E. cuniculi* and *A. locustae* genomes, the most significant is the shortening of intergenic regions. Interestingly, a comparison of ESTs from *A. locustae* and the two nucleomorph genomes revealed a correlation between the level of genome compaction and the frequency of overlapping transcripts [9]. At the extreme, this is a somewhat trivial observation, since transcripts could hardly be expected to terminate within the shortest intergenic regions, which can be as small as a few base pairs. However, it is still possible that the reduction of intergenic spaces leads to overlapping transcription in more interesting ways. To investigate this possibility, we compared the nature of transcripts from loci in different genomic contexts (i.e., more or less gene dense).

Overall, 65% of gene pairs harboring intergenic regions less than 150 bp were found to have overlapping transcripts in *A. locustae* (when including both 5′- and 3′-RACE products), corresponding to an almost sixfold increase over estimates based on ESTs [10]. When pairs of loci were separated by intergenic

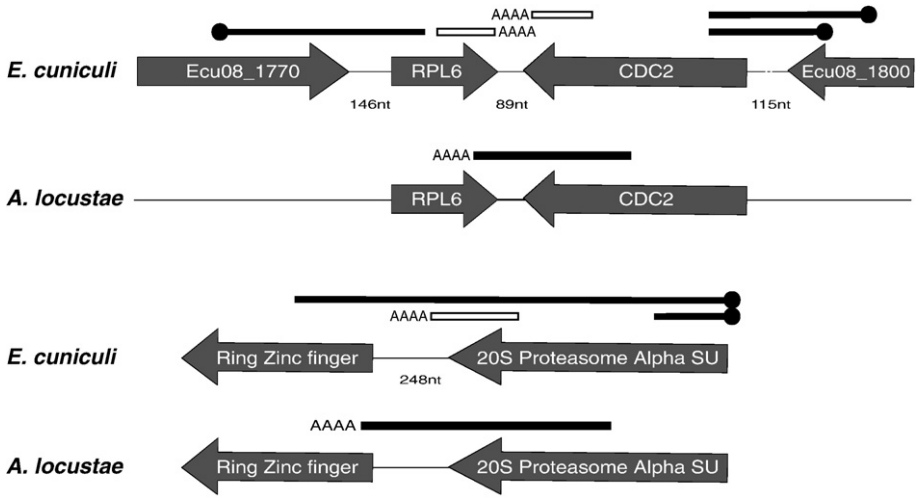

Fig. 1. Transcript starts and ends for adjacent genes for which multigene transcription was previously reported in *A. locustae* and that are in synteny with *E. cuniculi.* Gray arrows represent the position and direction of the genes in genomic DNA (with gene names inside). White rectangles represent the length and position of 3′-RACE fragments, with the poly(A) tail represented by four A's. Black rectangles represent the length and position of the 5′-RACE fragments, with the cap represented by a dot. The blunt part of the rectangles represents the location of the primer we used to perform the PCR. The length in nucleotides (nt) of the intergenic regions is shown. The length and position of the *A. locustae* EST fragments [10] are represented by the black rectangles and their respective poly(A) tails.

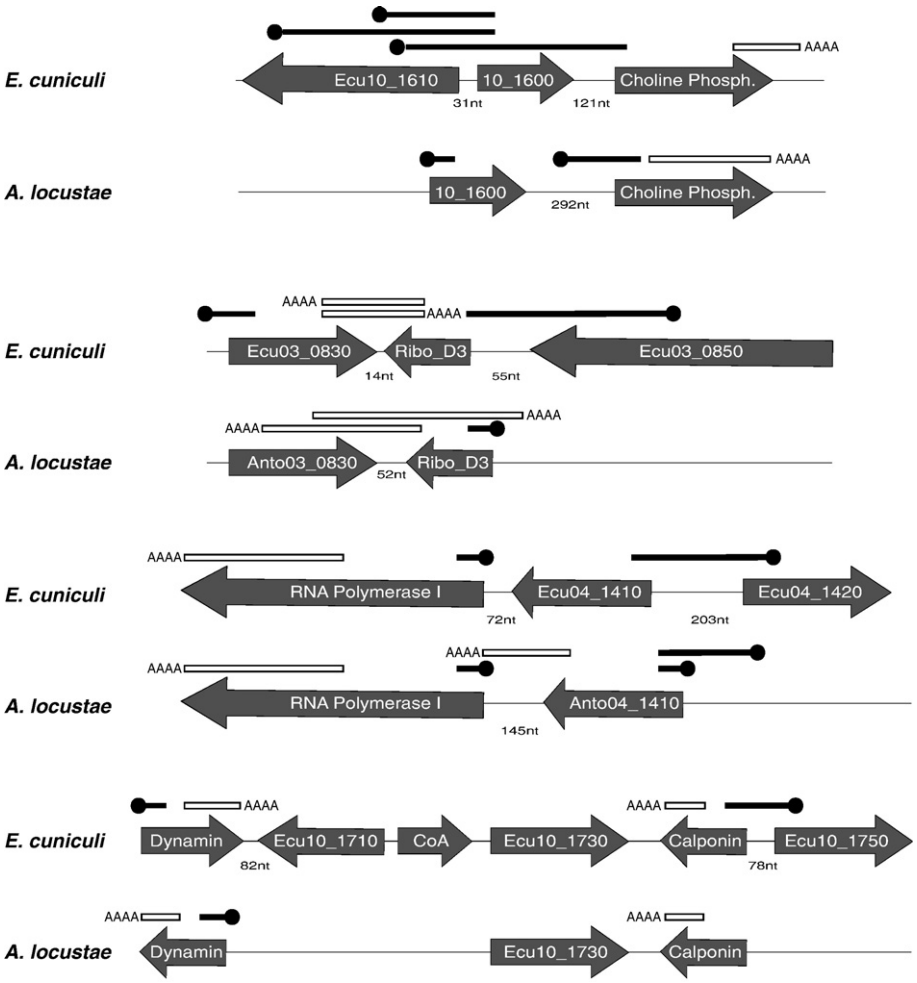

Fig. 2. Transcript start and end points for adjacent genes with short intergenic regions and in synteny between *E. cuniculi* and *A. locustae.* Gray arrows represent the position and direction of the genes in genomic DNA (with gene names inside). 5′- and 3′ RACE products are represented as in Fig. 1. The length in nucleotides (nt) of the intergenic regions is shown.

regions greater than 150 bp, the level of multigene transcription dropped to 10%. In *E. cuniculi* no such correlation can be made because, on average, the intergenic regions are much smaller than in *A. locustae.* However, overlapping transcription was found in more than 80% of the loci analyzed (when including both 5′- and 3′-RACE products). Whether the substantially higher frequency of overlap we observe in *E. cuniculi* is due to the loci sampled or to a real difference in transcription that is reflected across the genome is not entirely clear, but the very high frequency of overlap observed in *E. cuniculi* is not without precedent, since the frequency observed in the *Guillardia theta* nucleomorph genome is also close to 100% [10].

Differences in the lengths of intergenic regions are also a likely explanation for the absence of overlapping transcription in two *E. cuniculi* loci studied previously (spore wall protein and polar tube protein 3) [19,20]. These genes are separated from their neighbors by significantly larger than average intergenic regions (e.g., about 500 bp downstream for the spore wall protein and about 250 bp upstream in the case of polar tube protein 3). These distances are also greater than those analyzed here (which are 99 bp on average, close to the average for the genome). More interestingly, a 165-bp 3′ and a 172-bp 5′ UTR have been reported for mRNAs from these genes. If applied to our dataset, such long UTRs would have resulted in overlapping transcription in more than 80% of the cases.

In a few regions of the genome where two genes are encoded in a message in the same strand (i.e., hypothetical protein–actin, Vid25–ubiquitin, ring zinc finger–20S proteasome), current data are consistent with two possibilities: either one gene is subsequently translated while the other is simply an unexpressed UTR or the genes are coexpressed, as is the case for operons. Additional data would be required to distinguish between these possibilities; however, there are some observations that allow us to suggest the first possibility is more likely. Unlike these few cases, the majority of overlapping transcripts are not consistent with the presence of operons, either because the genes are encoded on opposite strands or because one gene is not entirely transcribed, so we know the first possibility is relatively common in these genomes. The presence of even a few operon-like transcription units would require characteristics of the ribosome not normally found in eukaryotes or, alternatively, a molecular system to process polycistronic mRNAs into individual messages, a system for which there currently are no data. Because frequent overlapping transcription would have

to be proposed along with operon-like cotranscription anyway, and because the presence of operon-like cotranscription would require additional mechanisms for which there is as yet no evidence, we currently favor the conclusion that none of these transcripts directs the expression of more than one gene.

A lack of conservation in transcription processing sites between microsporidian species

If transcription initiation and termination signals for a gene are located within adjacent genes, it could be difficult to separate those genes in the genome without deleterious effect, and overlapping transcription could thus contribute to genome stability. If this were the case, one might expect to find a high degree of conservation of the genes that have overlapping transcription and perhaps even the locations of end points between species in which gene order was also conserved. Alternatively, no correspondence between transcript overlap in the two species would indicate that this process is relatively fluid and that processing points change and adapt more quickly than the gene order.

The transcription profiles we observed between *A. locustae* and *E. cuniculi* were significantly different. In *A. locustae* the overall transcriptional starting points were conservative and consistently found within a few base pairs of the start codons, whereas transcripts terminated at various positions and often within downstream genes. In *E. cuniculi,* the pattern was completely reversed: a very high rate of overlap was detected at the 5′ ends, while termination points were rather consistent and generally intergenic. These species-specific patterns were maintained in cases in which the gene order was conserved between the two species, overall suggesting that transcription processing signals are evolving faster than gene order and, therefore, overlapping transcription is unlikely to be a significant force in the conservation of the genome architecture.

Variability of transcription in microsporidia

We have shown that the characteristic of overlapping transcription is common to two distantly related microsporidia and that the pattern of overlap in each genome is generally consistent and predictable, but is different between the two genomes. This is true even at loci where the order of genes is strictly conserved between the genomes, suggesting these differences reflect real variation in transcription control in these species. In directly comparable loci the pattern typical of the genome as a whole was a better predictor for the state of any given loci than the state of the homologous loci in another species, suggesting that the exact location of initiation and termination signals is evolving very quickly.

In practical terms, these observations present even more problems for the study of gene expression in microsporidia. When the high frequency of overlapping transcription was first characterized, it was noted that techniques such as SAGE (serial analysis of gene expression) and arrays that are widely used to study parasite transcriptomes [21–23] would be more difficult to apply to such a genome because the overlapping mRNAs would add an extra layer of complexity to the interpretation of a signal for any given gene. This is not to say that it would be impossible to interpret array data from a microsporidian, but if a small number of potential targets for a given gene were available, it would be difficult to know if binding to those targets was due to the expression of that gene or adjacent genes without prior knowledge of transcriptional overlap in that region of the genome. With increasing coverage of the genome on an array, signals could presumably be attributed to a particular gene with increasing confidence, but the interpretation would be significantly more complex than it is for genomes where transcription does not overlap to any great extent. The current data add to this problem in several ways. We show that the complexities of overlapping transcription extend to the human parasite, *E. cuniculi;* that the frequency of overlap in this species is even higher than that reported in *A. locustae;* and that the pattern of overlap is more complex than that of *A. locustae,* since it more often involves 5′ ends. Each of these factors can compound the complexity of the interpretation and also suggests that information on the frequency and nature of transcriptional overlap determined from one genome will not necessarily be applicable to other microsporidia.

Materials and methods

A. locustae and *E. cuniculi* spores were disrupted by beating with glass beads, and total RNA was extracted using an RNAqueous kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. We chose to use spores over intracellular stages to allow a direct comparison with previous data obtained by others [10,12]. 3′- and 5′-RACE was performed using the First Choice RLM-RACE kit (Ambion). For 5′-RACE, this protocol involves sequential treatments with calf intestinal phosphatase (CIP) and tobacco acid pyrophosphatase (TAP) prior to an RNA–RNA adaptor ligation with T4-RNA ligase. This selects for full-length, capped mRNAs. Fragments were amplified using gene-specific nested primers (see Supplemental Tables 1 and 2). Amplification products were cloned with TOPO TA vector (Invitrogen) and four clones of each sequenced. To ensure that all 5′-RACE fragments we obtained were strictly cap-dependent, and to avoid the analysis of contaminating genomic DNA and noncapped truncated cDNA fragments, 200 ng of CIP-treated total RNA (two times more than the amount usually required for the reaction) was subjected to the same conditions in the absence of TAP. All such controls yielded no product or products of completely different sizes compared to those obtained through the CIP–TAP procedure. In two cases both the CIP–TAP and "minus-TAP" yielded products of a similar size and these results were excluded from this study. In *A. locustae* 23 5′- and 23 3′-RACE fragments from 25 loci were analyzed. In *E. cuniculi* 19 5′- and 11 3′-RACE fragments from 16 loci were analyzed. Ideally both 5′- and 3′-RACE fragments were analyzed for a given locus, but in some cases only one could be amplified using various conditions and primers. When more than a single band was amplified through the PCR, all products were systematically cloned and sequenced to ensure they are derived from the expected region of the genome. In cases in which they do, their location and length are included in the figures and their sequence was deposited with GenBank. In a few cases a product was sequenced but discarded, for example, 3′-RACE sequences in which the poly(A) "tail" was found to correspond to an A track in the genome and 5′-RACE products that when sequenced were shown to derive from another region of the genome. In total, 41 loci were analyzed in this study and the 76 cDNA fragments corresponding to 5′ and 3′ transcriptional ends were deposited with GenBank under Accession Nos. ES880821 to ES880896 (Supplemental Tables 1, 2a, and 2b).

Acknowledgments

This work was supported by a grant from the Canadian Institutes for Health Research (MOP-84265). N.C. is supported by a fellowship from the Swiss National Science Foundation

(PBLAA–114238). P.J.K. is a fellow of the Canadian Institute for Advanced Research and a Senior Scholar of the Michael Smith Foundation for Health Research. We thank Rowena Stern, Lena Burri, and two anonymous reviewers for critically reading the manuscript and Naomi Fast and Renny Lee for providing *E. cuniculi* RNA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2007.12.006.

References

[1] J.I.R. Larsson, Identification of microsporidia, Acta Protozool. 38 (1999) 161–197.
[2] J.J. Becnel, T.G. Andreadis, Microsporidia in insects, in: W.L.M. Witter (Ed.), The Microsporidia and Microsporidiosis, Am. Soc. Microbiol. Press, Washington, DC, 1999, pp. 447–501.
[3] C.R. Vossbrinck, J.V. Maddox, S. Friedman, B.A. Debrunner-Vossbrinck, C.R. Woese, Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes, Nature 326 (1987) 411–414.
[4] P.J. Keeling, Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia, Fungal Genet. Biol. 38 (2003) 298–309.
[5] Y. Van De Peer, A. Ben Ali, A. Meyer, Microsporidia: accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi, Gene 246 (2000) 1–8.
[6] P.J. Keeling, W.F. Doolittle, Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family, Mol. Biol. Evol. 13 (1996) 1297–1305.
[7] C. Biderre, M. Pages, G. Metenier, D. David, J. Bata, G. Prensier, C.P. Vivares, On small genomes in eukaryotic organisms: molecular karyotypes of two microsporidian species (Protozoa) parasites of vertebrates, C. R. Acad. Sci. III 317 (1994) 399–404.
[8] E. Peyretaillade, C. Biderre, P. Peyret, F. Duffieux, G. Metenier, M. Gouy, B. Michot, C.P. Vivares, 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 (1998) 3513–3520.
[9] M.D. Katinka, S. Duprat, E. Cornillot, G. Metenier, F. Thomarat, G. Prensier, V. Barbe, E. Peyretaillade, P. Brottier, P. Wincker, F. Delvac, H. El Alaoui, P. Peyret, W. Saurin, M. Gouy, J. Weissenbach, C.P. Vivares, Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi, Nature 414 (2001) 450–453.
[10] B.A. Williams, C.H. Slamovits, N.J. Patron, N.M. Fast, P.J. Keeling, A high frequency of overlapping gene expression in compacted eukaryotic genomes, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 10936–10941.
[11] M.J. Jacob, Genetic regulatory mechanisms in the synthesis of proteins, J. Mol. Biol. 3 (1961) 318–356.
[12] C.H. Slamovits, P.J. Keeling, Class II photolyase in a microsporidian intracellular parasite, J. Mol. Biol. 341 (2004) 713–721.
[13] M. Gerads, J.F. Ernst, Overlapping coding regions and transcriptional units of two essential chromosomal genes (CCT8, TRP1) in the fungal pathogen Candida albicans, Nucleic Acids Res. 26 (1998) 5061–5066.
[14] K. Hansen, C.E. Birse, N.J. Proudfoot, Nascent transcription from the nmt1 and nmt2 genes of Schizosaccharomyces pombe overlaps neighbouring genes, EMBO J. 17 (1998) 3066–3077.
[15] F. Nilsen, C. Endresen, I. Hordvik, Molecular phylogeny of microsporidians with particular reference to species that infect the muscles of fish, J. Eukaryot. Microbiol. 45 (1998) 535–543.
[16] C.H. Slamovits, B.A. Williams, P.J. Keeling, Transfer of Nosema locustae (Microsporidia) to Antonospora locustae n. comb. based on molecular and ultrastructural data, J. Eukaryot. Microbiol. 51 (2004) 207–213.
[17] J. Xu, G. Pan, L. Fang, J. Li, X. Tian, T. Li, Z. Zhou, Z. Xiang, The varying microsporidian genome: existence of long-terminal repeat retrotransposon in domesticated silkworm parasite Nosema bombycis, Int. J. Parasitol. 36 (2006) 1049–1056.
[18] N. Corradi, D.E. Akiyoshi, H.G. Morrison, X. Feng, L.M. Weiss, S. Tzipori, P.J. Keeling, Patterns of genome evolution among the microsporidian parasites *Encephalitozoon cuniculi*, *Antonospora locustae* and *Enterocytozoon bieneusi*, PLoS ONE 2 (2007) e1277, doi:10.1371/journal.pone.0001277.
[19] W. Bohne, D.J. Ferguson, K. Kohler, U. Gross, Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian patogen Encephalitozoon cuniculi, Infect. Immun. 68 (2000) 2268–2275.
[20] I. Peuvel, P. Peyret, G. Metenier, C.P. Vivares, F. Delbac, The microsporidian polar tube: evidence for a third polar tube protein (PTP3) in Encephalitozoon cuniculi, Mol. Biochem. Parasitol. 122 (2002) 69–80.
[21] D.P. Knox, P.J. Skuce, SAGE and the quantitative analysis of gene expression in parasites, Trends Parasitol. 21 (2005) 322–326.
[22] P.T. Monis, S. Giglio, A.R. Keegan, R.C. Andrew Thompson, Emerging technologies for the detection and genetic characterization of protozoan parasites, Trends Parasitol. 21 (2005) 340–346.
[23] G.N. Gobert, L.P. Moertel, D.P. Mcmanus, Microarrays: new tools to unravel parasite transcriptomes, Parasitology 131 (2005) 439–448.