Review

On the origin of mitochondria and Rickettsia-related eukaryotic endosymbionts

B. Franz Lang[1*], Henner Brinkmann[1], Liisa B. Koski[1], Masahiro Fujishima[2], Hans-Dieter Görtz[3] and Gertraud Burger[1]

[1]*Program in Evolutionary Biology, Canadian Institute for Advanced Research; Centre Robert Cedergren, Département de Biochimie, Université de Montréal, 2900 Boulevard Edouard-Montpetit, Montréal, Québec, H3T 1J4, Canada.*
[2]*Biological Institute, Faculty of Science, Yamaguchi University, Yamaguchi 753-8512, Japan.*
[3]*Abteilung Zoologie, Biologisches Institut, Universität Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany.*

SUMMARY

Resent insights into the origin and early evolution of mitochondria come from two approaches: the investigation of mtDNAs from minimally derived (primitive) mitochondriate eukaryotes, in particular jakobid flagellates, and of genomes from intracellular α-proteobacterial symbionts. Of particular interest in this context is *Holospora obtusa*, an intracellular bacterial endosymbiont that resides and replicates in the somatic nucleus of its eukaryotic host, the ciliate *Paramecium caudatum*. Currently we have sequenced close to 50% of the ~ 1.7 Mbp *H. obtusa* genome, revealing the absence of genes for oxidative phosphorylation, the TCA cycle, and many other metabolic pathways, but the presence of several pathogenesis-related genes and a high number of bacterial IS elements. Phylogenetic analyses with multiple protein sequences place *H. obtusa* basally to the Rickettsia-Ehrlichia-Wolbachia assemblage of bacterial pathogens. This leads us to postulate that *H. obtusa* is the closest bacterial relative of mitochondria known to date.

*Corresponding author
E-mail: Franz.Lang@Umontreal.ca

Received: 28 June 2005.

INTRODUCTION

One of the major advancements in understanding eukaryotic evolution was the discovery that mitochondria evolved from an endosymbiotic α-Proteobacterium, and that mitochondrial DNA (mtDNA) is a relict bacterial genome. Most mtDNAs have lost earmarks of their bacterial origin, with a notable exception in jakobid flagellates

Table 1: Biological processes in mitochondria [1]

DNA replication, recombination and repair
Transcription[2] (a phage-like RNA polymerase, except in jakobid flagellates using an **$\alpha_2\beta\beta'\sigma$ polymerase[3]**)
RNA processing (including RNA editing and intron splicing)
Translation (rRNAs, tRNAs and ribosomal proteins, **EFTU[3]** and **tmRNAs[3]**)
Protein and RNA translocation (Tom, Tim, Tob, **Tat (*tatA*[3]**, *tatC*), **Sec[3]**, Oxa1, and VDAC translocases)
Mitochondrial carriers (translocators of ADP/ATP, co-factors, metabolites, substrates, etc.)
Chaperones
Mitochondrial morphology and division
Proteases
Nucleases
Electron transport and ATP synthesis
Lipid metabolism
Nucleotide metabolism
Amino acid metabolism
Carbohydrate metabolism
Fe-S biosynthesis and export
Heme biosynthesis
Ubiquinone biosynthesis
Enzyme cofactor biosynthesis

[1] Note that some pathways serve mitochondria-specific functions only, whereas others supply the whole cell.
[2] Processes in bold involve mtDNA-encoded components.
[3] MtDNA-encoded only in *Reclinomonas americana* (Lang et al. 1997; Jacob et al. 2004).
For a compilation of genes from completely sequenced mtDNAs see GOBASE at http://amoebidia.bcm.umontreal.ca/pg-gobase/searches/compilations.php

(Lang et al. 1997; Jacob et al. 2004). According to current estimates, mitochondria were acquired more than one billion years ago, which makes inferences of early steps in mitochondrial evolution a non-trivial problem. Not surprisingly, the nature of the partners implicated in mitochondrial endosymbiosis, the driving forces and timing of their fusion, and initial stages of reductive mitochondrial genome evolution, remain vague.

Contemporary mtDNAs have retained only 5 – 100 from the original set of several hundred to several thousand genes residing in bacterial genomes. One portion of these genes was lost in the proto-mitochondrial genome, another one has migrated to the nucleus during early steps of endosymbiosis. The total number of nuclear genes that are involved in mitochondrial biogenesis has been estimated at 700-800 proteins in yeast (see Table 1 for biological processes in which mitochondria of yeast and other eukaryotes are involved). This estimate is based on mass spectrometry of proteins from isolated mitochondria (Sickmann et al. 2003) and sub-cellular localization studies (Kumar et al. 2002; Huh et al. 2003), combined with genomic approaches (Prokisch et al. 2004). However, these numbers come with a high degree of imprecision (Reichert and Neupert 2004). In addition, only a small fraction of yeast mitochondrial proteins can

be placed with confidence in phylogenetic analyses; consequently, it remains unknown how many of the nucleus-encoded mitochondrial proteins derive from the α-proteobacterial ancestor, the eukaryotic host, or other sources (Burger and Lang 2003).

To address the question of mitochondrial origins, we have begun to investigate the genome of the bacterium *Holospora obtusa*, an endosymbiont of the ciliate *Paramecium caudatum* (Görtz et al. 1990). The objective of our study is to conduct genome comparisons between intracellular *versus* free-living bacteria on one hand, and endosymbiotic bacteria *versus* organelles on the other.

In the present review, we will give a brief update on the diversity of mitochondrial genomes across eukaryotes and discuss the early evolutionary events that led to the establishment of mitochondria. Then, we will present a preliminary analysis of the *Holospora obtusa* genome sequence, and conclude with a phylogenetic analysis of this species.

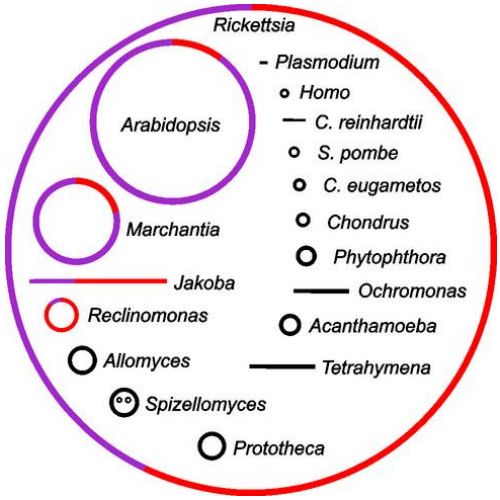

Figure 1: Genome size and coding content of mitochondria and R*ickettsia*. Circles and lines represent circular and linear genome shapes, respectively. For genomes >60 kbp, the portion coding for genes with known function (red) is distinguished from that including intergenic regions and ORFs (purple).

Distribution of mitochondria across eukaryotes

Not all eukaryotes contain mitochondria. Exceptions, although rare, are species that live in anoxic environments, such as *Trichomonas* (parabasalids), *Giardia* and *Hexamita* (diplomonads), *Entamoeba* and *Mastigamoeba* (pelobionts), *Retortamonas* (retortamonads), *Microsporidia* (fungi) and *Neocallimastix* (fungi). Several of these species have most likely lost functional mitochondria secondarily rather than being primitively amitochondriate, because they contain nucleus-encoded genes that are most likely of mitochondrial origin (e.g., the ADP/ATP carrier (van der Giezen et al. 2002; Voncken et al. 2002) and the 60 and 70 kDa heat-shock proteins (van der Giezen et al. 2003)). In addition, a growing number of amitochondriate eukaryotes are reported to contain organelles (hydrogenosomes, cryptosomes or mitosomes), which are believed to be derived from mitochondria because they import the proteins mentioned above (Williams and Keeling 2003). Taken together, this raises the possibility that eukaryotes without mitochondria may never have existed. Alternatively, their descendants might have not survived the increase of oxygen concentration on Earth that was due to the emergence and proliferation of photosynthetic eukaryotes.

Typical features of mitochondrial genomes – a brief overview

The more mitochondrial genetic systems are analyzed, the less holds the concept of a 'typical' mtDNA. Even so, certain features are frequently found across a broad taxonomical range of eukaryotes, as briefly summarized in the following (for more details and references, see (Gray et al. 1998; Lang et al. 1999; Gray et al. 2004)).

A+T content. With few exceptions, mitochondrial genomes feature a high A+T content (60-80%, compared to nuclear) a property often used for

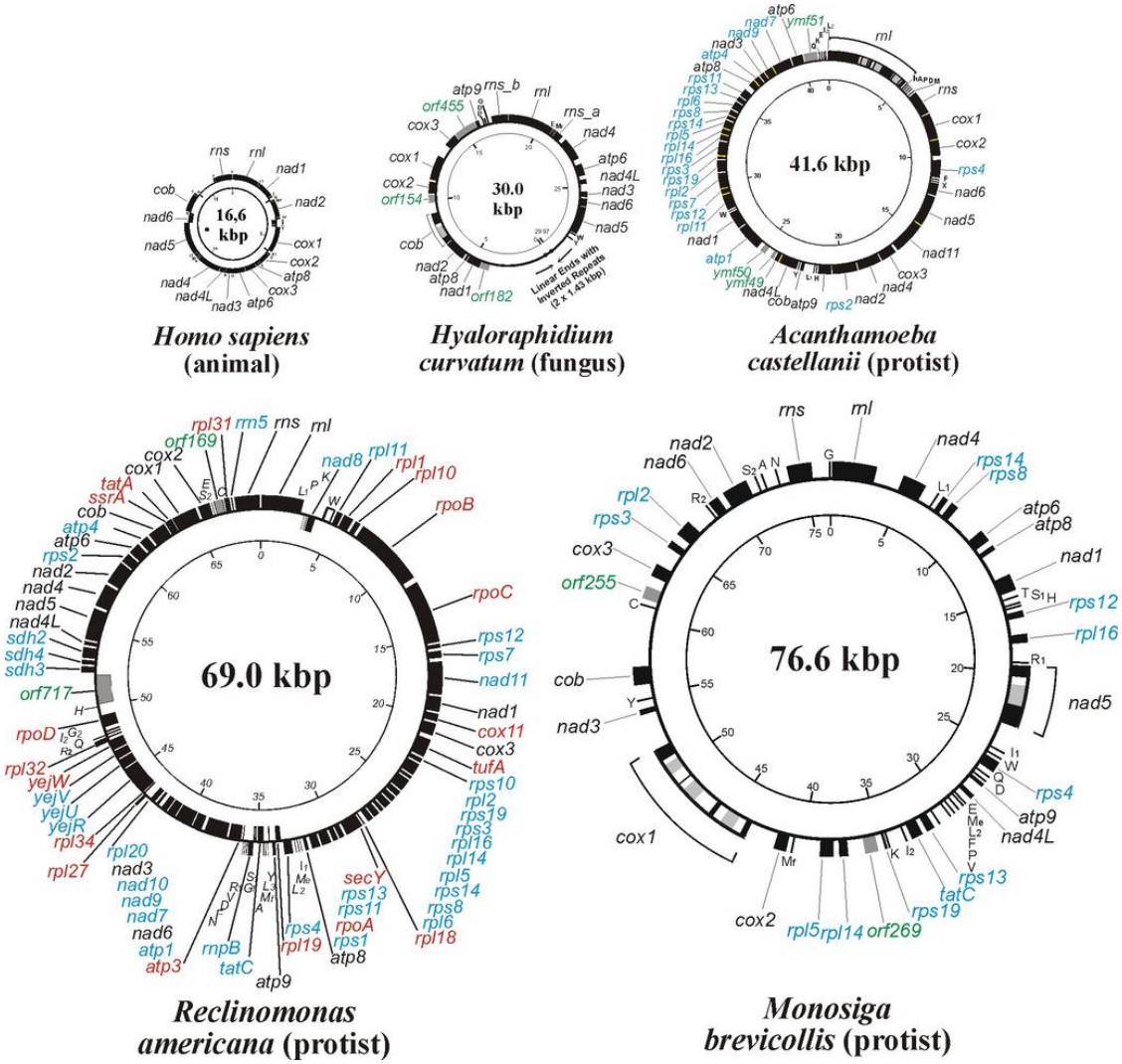

Figure 2: Mitochondrial genomes from diverse eukaryotic groups. Genetic maps of mtDNAs (see also (O'Brien et al. 2003) and http://amoebidia.bcm.umontreal.ca/pg-gobase/searches/map.php) from an animal (Anderson et al. 1981), a fungus (Forget et al. 2002) and three protists (Burger et al. 1995; Lang et al. 1997; Lang et al. 2002), demonstrating the wide range of coding capacity and organization (the mtDNA of *Hyaloraphidium* although linear, is depicted as a circle to facilitate comparisons). Black rectangles, identified genes; grey rectangles, ORFs; white rectangles with grey-shaded sections, introns and intronic open reading frames; rectangles on outer/inner ring, transcription is clockwise/counterclockwise. Gene names in black belong to the basic gene complement common to animal mtDNAs; green, ORFs; blue, extra genes present in various protists and plants; red, extra genes specific to jakobid flagellates.

physically separating mtDNA from other DNAs within an organism.

Size. The medium size of mtDNA is ~50 kbp, which corresponds to ~1% of the typical size of an α-proteobacterial genome from which it is derived. Yet deviations from the medium size by an order of magnitude are quite common (Fig. 1). With only 6 kbp, the mtDNA of the primate and bird parasite *Plasmodium* is the smallest known, while the largest ones occur in land plants, measuring from 180 to 2,400 kbp (Ward et al. 1981).

Shape. MtDNA molecules (referred to in the following as chromosomes) come in three major topologies (for a review, see (Bendich 1996)). In most organisms, mitochondrial chromosomes are circular-mapping but linear in structure, made up of several tandemly concatenated copies. True (monomeric) circular or linear mtDNAs have been shown to occur in only a limited number of phylogenetically disparate taxa such as fungi, vertebrates and Euglenozoa (e.g., (Forget et al. 2002; Nosek and Tomaska 2003; Marande et al. 2005) and references therein).

Chromosome number and ploidy. Mitochondrial DNA typically consists of a single kind of molecule in numerous copies, ranging from ~10 to ~1000 per cell, i.e., a dozen or more per organelle and often more than one organelle per cell.

Gene content. Invariantly, mtDNA-encoded components participate in processes that take place in mitochondria, including respiration, oxidative phosphorylation and translation; less frequently in protein import/maturation and RNA processing, and rarely in transcription (Table 1, Fig. 2). Again, at the extremes are *Plasmodium* and its relatives with as few as five genes (three of which code for proteins and two for rRNAs), and jakobid flagellates with nearly 100 genes (including 70 proteins and 25-26 tRNAs). The number of mtDNA-encoded tRNAs is most variable across eukaryotes (0 to >30). The reason for the up to 20-fold difference in mitochondrial gene content is variation in the extent of gene migration to the nucleus (Adams and Palmer 2003; Burger and Lang 2003).

The most bacteria-like mtDNA is found in *Reclinomonas americana*

In past years, the eubacterial ancestry of mitochondria has become strikingly obvious through the study of minimally derived protists, such as the recently recognized jakobid flagellate *Reclinomonas americana* (Lang et al. 1997) (Fig. 2). Among the most eubacteria-like features of jakobid mtDNAs are putative ribosome binding motifs (Shine-Dalgarno sequences) upstream of protein-coding genes, multi-subunit RNA polymerase, Sec-dependent protein import, transfer messenger RNA (tmRNA) involved in the release of stalled ribosomes, and the elongation factor Tu (Lang et al. 1997; Jacob et al. 2004). Neither of these has been found so far in mtDNAs of other eukaryotes.

Mitochondrial *versus* bacterial genomes

Genomes of free-living bacteria usually contain several thousand genes, compared to only ~ 500 to 1000 in endosymbiotic bacteria. For example, in the obligatory pathogen *Haemophilus influenza*, the causative agent of meningitis, 636 genes could be assigned to a function and a further 363 are predicted to code for proteins of unknown function (TIGR http://pohh.phy.ncu.edi/tw/cdy/USS/). The gene content of mitochondrial genomes is yet an order of magnitude smaller than those of pathogenic/endosymbiotic bacteria. Completely absent from mtDNAs are genes involved in several basic cellular processes such as DNA replication, recombination, motility, metabolism, and signal transduction (Fig. 2, Table 1).

As discussed in detail elsewhere (see (Burger and Lang 2003) and references therein), gene loss and elevated genomic A+T content in bacteria appear to progress concurrently with the transition from a free-living to an intracellular lifestyle. Interestingly, the trends of genome evolution associ-

ated with intracellularity are shared by bacterial pathogens of different phylogenetic affiliation: α-Proteobacteria (*Rickettsia*, *Wolbachia, Ehrlichia*), β-Proteobacteria (*Neisseria*), γ-Proteobacteria (*Haemophilus*, *Buchnera*, *Wigglesworthia*), Firmicutes (*Mycoplasma*), the Chlamydia group, and Spirochetes (*Borrelia*). Obviously, by adapting to intracellular environments, these organisms have undergone convergent evolution. A similar convergence is found in mitochondria and plastids.

Do mitochondria originate from endosymbiotic or free-living α-Proteobacteria?

Phylogenetic analyses provide clear evidence that mitochondria originated from within α-Proteobacteria (e.g., (Andersson et al. 1998; Burger et al. 1999)), a bacterial assemblage that includes free-living species such as *Magnetospirillum* and *Rhodobacter*, facultative symbionts/pathogens of plants such as *Rhizobium* and *Agrobacter*, and intracellular obligatory animal pathogens of the *Rickettsia* group including the causative agents of typhus and spotted fever. However, phylogenetic reconstructions, even when including mitochondrial data from the minimally derived jakobids, do not resolve with confidence which of these α-proteobacterial groups is most closely related to mitochondria. To address this question with a broader dataset, we began sequencing the genome of the α-proteobacterial *Holospora obtusa,* an intracellular endosymbiont of the ciliate *Paramecium caudatum*. In the following, we briefly describe this bacterium, its life style and genome, before going on to revisit the ancestry of mitochondria using these new data.

RESULTS AND DISCUSSION

The highly reduced genome of the endosymbiotic bacterium *Holospora obtusa*

Ciliates harbor a variety of bacteria that enter their host *via* the food vacuole. Some of these bacterial species further invade the cytoplasm as well as the micro- and the macronucleus where they reside and reproduce. About 200 ciliate species have been recorded to host intracellular bacteria; the *Paramecium* genus alone harbors close to 60 different bacterial taxa (reviewed in (Fokin et al. 2004)).

The specific host of *Holospora obtusa* is the ciliate *Paramecium caudatum*, on which the bacterium depends for propagation. Previous studies have uncovered the complex life cycle of *Holospora*, including the invasion of the somatic (macro-) nucleus and its interactions with and modulation of the host's subcellular structures upon infection (Görtz et al. 1990). In the course of transition from the infectious to the reproductive bacterial form, patterns of protein expression change considerably. So far, more than 20 bacterial proteins that are specifically involved in infection have been characterized by immunological methods and N-terminal microsequencing (e.g., (Görtz et al. 1990; Wiemann and Görtz 1991; Dohra et al. 1997; Dohra et al. 1998; Nakamura et al. 2004); Fujishima and Görtz, unpublished results).

We have decided to sequence the complete genome of *Holospora* for two reasons, (i) to facilitate identification of genes implicated in infection and (ii) to revisit the question about the phylogenetic position of this endosymbiont. At the time of writing, we have sequenced close to 50% of the 1.7 mbp genome size estimated by Timofeyeva AS, Rautian M and Görtz HD (unpublished). With an A+T-content of ~ 65%, the *Holospora* genome size is only moderately A+T rich, compared to other endosymbiotic bacteria, which have A+T contents up to 75% (reviewed in (Wernegreen 2002)).

As of now, more than 353 genes have been identified in the *Holospora* genome, including

~200 genes of known function (coding for 181 distinct proteins and 18 structural RNAs). The remaining ~150 hypothetical proteins (ORFs) have no recognizable counterparts in other organisms. The majority of assigned genes participate in basic cellular processes (replication and repair, transcription and translation, ATP synthesis), a wide variety of transport functions (ABC transporters) and a few metabolic pathways (e.g., lipid, heme, glucan, nucleotide and alanine biosynthesis). Most pathways considered essential for autotrophy appear to be completely missing, except, curiously, biotin synthesis (note that biotin synthesis is lacking in *Rickettsia*). We speculate that the provision of biotin constitutes one of the benefits that *Paramecium* receives from its association with *Holospora*. Other potential benefits for the host cell are longer survival of infected *Paramecium* at low temperatures (10 ºC) and longer survival and increased cell motility at 37 ºC, which has been associated with over-expression of Hsp70 and/or GroEL (Hori and Fujishima 2003; Fujishima et al. 2005).

Intriguingly, we found in the *Holospora* genome no genes for respiratory chain complexes, which in aerobic organisms serve to drive efficient ATP synthesis *via* oxidative phosphorylation. It appears that *Holospora* relies on 'energy parasitism' by ATP import *via* its ADP/ATP transporter system (Linka et al. 2003). This feature distinguishes *Holospora* from *Rickettsia* and its relatives who have retained the oxidative phosphorylation system (e.g., (Andersson et al. 1998; Renesto et al. 2005)), as well as the TCA cycle and the glycolysis III pathway (pyruvate → acetate). The loss of respiration and oxidative phosphorylation in *Holospora* is not surprising, given that this endosymbiont resides and reproduces in the host's nucleus where oxygen supply is limited.

We have discovered additional genes in the *Holospora* genome whose products play a role in host invasion and interact with the host's subcellular components (Fujishima M, Görtz HD, Lang BF and Burger G, unpublished). These genes code for components of the general secretion pathway, lipoprotein carriers, and permeases. The *Holospora* genome also contains a class of ~ 40 nearly identical insertion elements (classified as transposase_11; pfam01609) that are most similar to counterparts in *Legionella pneumophila* ((Chien et al. 2004); locus CAH17112)) and *Francisella tularensis* ((Larsson et al. 2005); locus YP_169173)). Finally, we identified numerous other repeat elements, some of which have similarities to transposase-containing IS sequences of bacteria, and others that appear to be *Holospora*-specific.

Holospora obtusa is only distantly related to Rickettsiales

Initial phylogenetic analyses with small subunit rRNA (SSU-rRNA) sequences affiliated *Holospora obtusa* with *Caedibacter caryophila* (another *Paramecium* endosymbiont; (Springer et al. 1993)), which together emerge from within the Rickettsia lineage of α-Proteobacteria (e.g., (Springer et al. 1993; Gray and Spencer 1996; Horn et al. 1999)). However, these results have to be met with reservations because the statistical support for this placement with SSU-rRNA data is non-significant (< 95% bootstrap support; likelihood ratio tests were not performed). In addition, the clustering of *Rickettsia*- and *Holospora*-relatives may be solely due to their elevated A+T content, contrasting with the low A+T content of free-living α-Proteobacteria. Nucleotide sequence datasets of highly unbalanced A+T content are known to be prone to a strong phylogenetic artifact, whose effect is similar to that caused by unequal evolutionary rates known as long-branch attraction (Felsenstein 1978). In fact, long-branch attraction is itself a major concern, as all symbiotic/pathogenic bacteria have highly accelerated evolutionary rates as a consequence of their adaptation to an intracellular life style. In combination,

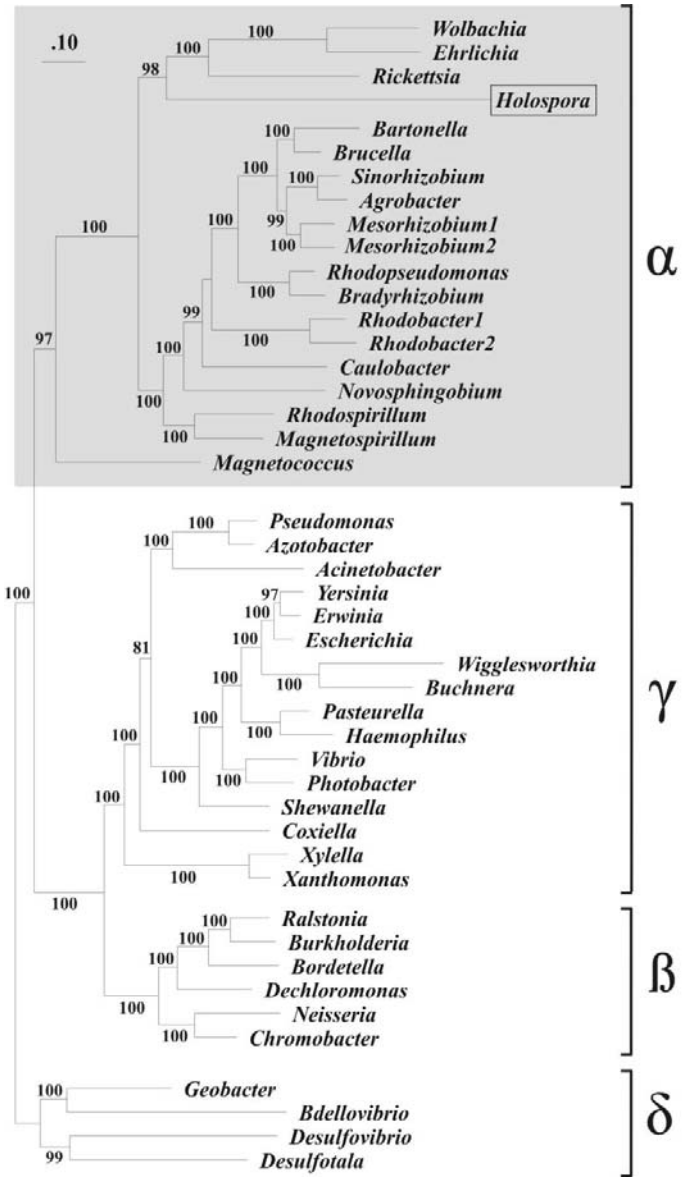

Figure 3: Phylogenetic relationship of *Holospora* to other α-Proteobacteria
The analysis is based on a concatenated dataset of bacterial proteins. The tree has been inferred using PhyML and the WAG+F+Γ model. Numbers at branches represent support values obtained with 100 bootstrap replicates. The scale bar denotes the estimated number of amino acid substitutions per site. The four eubacterial subdivisions (α, β, γ, δ) are indicated. For more information on the selection of sequences and the phylogenetic methodology used, see (Rodríguez-Ezpeleta et al. 2005).

the effect of A+T bias plus long-branch attraction might be strong enough to regroup *Holospora* with Rickettsiales in a positively misleading way.

Under such circumstances, it is advisable to employ phylogenetic approaches that are least sensitive to long-branch attraction (such as likelihood methods; (Delsuc et al. 2005)), and to use deduced protein rather than nucleotide sequences, in order to reduce the impact of A+T bias. Moreover, the analysis of multiple protein sequences helps to overcome the problem of non-significant statistical support, as does the application of the most realistic evolutionary models (e.g., an appropriate amino acid substitution matrix; among-site variation of evolutionary rates modeled by a gamma distribution).

Lateral (or horizontal) gene transfer is another potential problem in bacterial phylogeny

(Zhaxybayeva et al. 2004). Yet, a core of genes that generally code for essential proteins, and that are subunits of large complexes, are less likely affected by transfers (Brochier et al. 2002). These are the genes that we chose for the present analysis.

Fig. 3 shows a phylogenetic tree based on 6,109 amino acid positions from 18 genes. The proteobacterial relationships are resolved with significant support at all branches (except for a short internal branch in the γ subdivision). The impact of LGT seems negligible, because in multi-protein phylogenies, tree resolution and support values are expected to disappear with increasing horizontal gene transfer. In contrast to the results obtained with SSU-rRNA data mentioned above, *H. obtusa* branches in our protein-based analysis clearly outside the Rickettsia-Wolbachia-Ehrlichia lineage of bacterial pathogens.

Divergence point of mitochondria within α-Proteobacteria

Several studies by others suggest that mitochondria are specifically associated with *Rickettsiales*, or even emerged from within the Rickettsia lineage (e.g., (Gray and Spencer 1996; Andersson et al. 1998; Sicheritz-Ponten et al. 1998)). These conjectures need to be considered with caution because the corresponding analyses suffer from non-significant statistical support due to limited sequence data, or insufficient taxon sampling. We previously inferred phylogenies based on an extended set of mtDNA-encoded proteins (up to 13) that are highly conserved and present in most eukaryotic lineages, providing a total of ~ 3,000 amino acid positions (e.g., (Forget et al. 2002; Lang et al. 2002; Bullerwell et al. 2003a; Bullerwell et al. 2003b)). When including in this data set sequences from several α-proteobacteria but not *Holospora* (because it lacks the corresponding proteins), mitochondria branch between *Rickettsia* and *Sinorhizobium/Magnetospirillum* and somewhat closer to the latter two species than to *Rickettsia.* Superimposition of the bacterial tree shown in Fig. 3 with our bacterial/mitochondrial trees predicts that *H. obtusa* is closer to the mitochondrial divergence point than all other bacterial pathogens.

To test this prediction, a gene set will be required which is shared by *H. obtusa* and mitochondria, for instance nuclear-encoded mitochondrial proteins with counterparts in *Holospora*, such as ribosomal proteins, translation factors and chaperons. Work is in progress by the Canada-wide collaborative Protist EST Program (PEP; http://amoebidia.bcm.umontreal.ca/public/pepdb/agrm.php), to determine the expressed portion of nuclear genomes from a broad range of protists. These data will be instrumental for resolving the question whether *Holospora* is indeed the closest bacterial relative of mitochondria.

CONCLUSIONS

Our preliminary analyses of the *H. obtusa* genome show that α-proteobacterial endosymbionts represent an unexpected broad diversity, both phylogenetically and metabolically. The complete genome sequence of *H. obtusa* and its relatives such as *Caedibacter* will not only be key to better understanding eukaryotic host – pathogen interactions. These bacterial genome data, in combination with EST data from minimally derived eukaryotes, also promise to provide insight into the early history of mitochondrial endosymbiosis.

ACKNOWLEDGEMENTS

We thank M. Nickel, Y. Jacob and D. Tremblay for participation in *H. obtusa* genome sequencing, and L. Forget and Z. Wang for excellent technical assistance. This project was supported by grants from the Canadian Institutes for Health Research (MOP-42475, GX-15331), and by Genome Quebec and Genome Canada. Salary and

interaction support from the Canadian Institute for Advanced Research (BFL and GB) is gratefully acknowledged.

REFERENCES

Adams, K. L. and Palmer, J. D. (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phylogenet Evol 29, 380-395.

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R. and Young, I. G. (1981) Sequence and organization of the human mitochondrial genome. Nature 290, 457-465.

Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Ponten, T., Alsmark, U. C., Podowski, R. M., Naslund, A. K., Eriksson, A. S., Winkler, H. H. and Kurland, C. G. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396, 133-140.

Bendich, A. J. (1996) Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. J Mol Biol 255, 564-588.

Brochier, C., Bapteste, E., Moreira, D. and Philippe, H. (2002) Eubacterial phylogeny based on translational apparatus proteins. Trends Genet 18, 1-5.

Bullerwell, C. E., Forget, L. and Lang, B. F. (2003a) Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. Nucleic Acids Res 31, 1614-1623.

Bullerwell, C. E., Leigh, J., Seif, E., Longcore, J. E. and Lang, B. F. (2003b) Evolution of the fungi and their mitochondrial genomes. Pp. 133-160 *in* D. Arora, and G. G. Khachatourians, eds. Applied Mycology and Biotechnology. Elsevier Science, Amsterdam.

Burger, G. and Lang, B. F. (2003) Parallels in genome evolution in mitochondria and bacterial symbionts. IUBMB Life 55, 205-212.

Burger, G., Plante, I., Lonergan, K. M. and Gray, M. W. (1995) The mitochondrial DNA of the amoeboid protozoon, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. J Mol Biol 245, 522-537.

Burger, G., Saint-Louis, D., Gray, M. W. and Lang, B. F. (1999) Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea*. Cyanobacterial introns and shared ancestry of red and green algae. Plant Cell 11, 1675-1694.

Chien, M., Morozova, I., Shi, S., Sheng, H., Chen, J., Gomez, S. M., Asamani, G., Hill, K., Nuara, J., Feder, M., Rineer, J., Greenberg, J. J., Steshenko, V., Park, S. H., Zhao, B., Teplitskaya, E., Edwards, J. R., Pampou, S., Georghiou, A., Chou, I. C., Iannuccilli, W., Ulz, M. E., Kim, D. H., Geringer-Sameth, A., Goldsberry, C., Morozov, P., Fischer, S. G., Segal, G., Qu, X., Rzhetsky, A., Zhang, P., Cayanis, E., De Jong, P. J., Ju, J., Kalachikov, S., Shuman, H. A. and Russo, J. J. (2004) The genomic sequence of the accidental pathogen *Legionella pneumophila*. Science 305, 1966-1968.

Delsuc, F., Brinkmann, H. and Philippe, H. (2005) Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet 6, 361-375.

Dohra, H., Fujishima, M. and Ishikawa, H. (1998) Structure and expression of a GroE-homologous operon of a macronucleus-specific symbiont *Holospora obtusa* of the ciliate *Paramecium caudatum*. J Eukaryot Microbiol 45, 71-79.

Dohra, H., Yamamoto, K., Fujishima, M. and Ishikawa, H. (1997) Cloning and sequencing of gene coding for a periplasmic 5.4 kDa peptide

of the macronucleus-specific symbiont *Holospora obtusa* of the ciliate *Paramecium caudatum*. Zoolog Sci 14, 69-75.

Felsenstein, J. (1978) Cases in which parsimony and compatibility methods will be positively misleading. Syst Zool 27, 27-33.

Fokin, S. O., Skovorodkin, I. N., Schweikert, M. and Görtz, H. D. (2004) Co-infection of the macronucleus of *Paramecium caudatum* by free-living bacteria together with the infectious *Holospora obtusa*. J Eukaryot Microbiol 51, 417-424.

Forget, L., Ustinova, J., Wang, Z., Huss, V. A. and Lang, B. F. (2002) *Hyaloraphidium curvatum*: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. Mol Biol Evol 19, 310-319.

Fujishima, M., Kawai, M. and Yamamoto, R. (2005) *Paramecium caudatum* acquires heat-shock resistance in ciliary movement by infection with the endonuclear symbiotic bacterium *Holospora obtusa*. FEMS Microbiol Lett 243, 101-105.

Görtz, H. D., Lellig, S., Miosga, O. and Wiemann, M. (1990) Changes in fine structure and polypeptide pattern during development of *Holospora obtusa*, a bacterium infecting the macronucleus of *Paramecium caudatum*. J Bacteriol 172, 5664-5669.

Gray, M. and Spencer, D. (1996) Organellar evolution. Pp. 109-126 *in* D. Roberts, P. Sharp, G. Alderson, and M. Collins, eds. Evolution of Microbial Life. Cambridge University Press.

Gray, M. W., Lang, B. F. and Burger, G. (2004) Mitochondria of protists. Annu Rev Genet 38, 477-524.

Gray, M. W., Lang, B. F., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T. G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, Y. and Burger, G. (1998) Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 26, 865-878.

Hori, M. and Fujishima, M. (2003) The endosymbiotic bacterium *Holospora obtusa* enhances heat-shock gene expression of the host *Paramecium caudatum*. J Eukaryot Microbiol 50, 293-298.

Horn, M., Fritsche, T. R., Gautom, R. K., Schleifer, K. H. and Wagner, M. (1999) Novel bacterial endosymbionts of *Acanthamoeba spp*. related to the *Paramecium caudatum* symbiont *Caedibacter caryophilus*. Environ Microbiol 1, 357-367.

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. and O'Shea, E. K. (2003) Global analysis of protein localization in budding yeast. Nature 425, 686-691.

Jacob, Y., Seif, E., Paquet, P. O. and Lang, B. F. (2004) Loss of the mRNA-like region in mitochondrial tmRNAs of jakobids. RNA 10, 605-614.

Kumar, A., Agarwal, S., Heyman, J. A., Matson, S., Heidtman, M., Piccirillo, S., Umansky, L., Drawid, A., R., J., Liu, Y., Cheung, K. H., Miller, P., Gerstein, M., Roeder, G. S. and Snyder, M. (2002) Subcellular localization of the yeast proteome. Genes Dev 16, 707-719.

Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M. and Gray, M. W. (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387, 493-497.

Lang, B. F., Gray, M. W. and Burger, G. (1999) Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33, 351-397.

Lang, B. F., O'Kelly, C., Nerad, T., Gray, M. W. and Burger, G. (2002) The closest unicellular relatives of animals. Curr Biol 12, 1773-1778.

Larsson, P., Oyston, P. C., Chain, P., Chu, M. C., Duffield, M., Fuxelius, H. H., Garcia, E.,

Halltorp, G., Johansson, D., Isherwood, K. E., Karp, P. D., Larsson, E., Liu, Y., Michell, S., Prior, J., Prior, R., Malfatti, S., Sjostedt, A., Svensson, K., Thompson, N., Vergez, L., Wagg, J. K., Wren, B. W., Lindler, L. E., Andersson, S. G., Forsman, M. and Titball, R. W. (2005) The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. Nat Genet 37, 153-159.

Linka, N., Hurka, H., Lang, B. F., Burger, G., Winkler, H. H., Stamme, C., Urbany, C., Seil, I., Kusch, J. and Neuhaus, H. E. (2003) Phylogenetic relationships of non-mitochondrial nucleotide transport proteins in bacteria and eukaryotes. Gene 306, 27-35.

Marande, W., Lukes, J. and Burger, G. (2005) Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. Eukaryot Cell 4, 1137-1146.

Nakamura, Y., Aki, M., Aikawa, T., Hori, M. and Fujishima, M. (2004) Differences in gene expression of the ciliate *Paramecium caudatum* caused by endonuclear symbiosis with *Holospora obtusa*, revealed using differential display reverse transcribed PCR. FEMS Microbiol Lett 240, 209-213.

Nosek, J. and Tomaska, L. (2003) Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. Curr Genet 44, 73-84.

O'Brien, E. A., Badidi, E., Barbasiewicz, A., deSousa, C., Lang, B. F. and Burger, G. (2003) GOBASE--a database of mitochondrial and chloroplast information. Nucleic Acids Res 31, 176-178.

Prokisch, H., Scharfe, C., Camp, D. G., 2nd, Xiao, W., David, L., Andreoli, C., Monroe, M. E., Moore, R. J., Gritsenko, M. A., Kozany, C., Hixson, K. K., Mottaz, H. M., Zischka, H., Ueffing, M., Herman, Z. S., Davis, R. W., Meitinger, T., Oefner, P. J., Smith, R. D. and Steinmetz, L. M. (2004) Integrative analysis of the mitochondrial proteome in yeast. PLoS Biol 2, e160.

Reichert, A. S. and Neupert, W. (2004) Mitochondriomics or what makes us breathe. Trends Genet 20, 555-562.

Renesto, P., Ogata, H., Audic, S., Claverie, J. M. and Raoult, D. (2005) Some lessons from *Rickettsia* genomics. FEMS Microbiol Rev 29, 99-117.

Rodríguez-Ezpeleta, N., Brinkmann, H., Burey, S. C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H. J., Philippe, H. and Lang, B. F. (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae and glaucophytes. Curr Biol 15, 1325-1330.

Sicheritz-Ponten, T., Kurland, C. G. and Andersson, S. G. (1998) A phylogenetic analysis of the cytochrome b and cytochrome c oxidase I genes supports an origin of mitochondria from within the Rickettsiaceae. Biochim Biophys Acta 1365, 545-551.

Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H. E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B., Rehling, P., Pfanner, N. and Meisinger, C. (2003) The proteome of *Saccharomyces cerevisiae* mitochondria. Proc Natl Acad Sci U S A 100, 13207-13212.

Springer, N., Ludwig, W., Amann, R., Schmidt, H. J., Görtz, H. D. and Schleifer, K. H. (1993) Occurrence of fragmented 16S rRNA in an obligate bacterial endosymbiont of *Paramecium caudatum*. Proc Natl Acad Sci U S A 90, 9892-9895.

van der Giezen, M., Birdsey, G. M., Horner, D. S., Lucocq, J., Dyal, P. L., Benchimol, M., Danpure, C. J. and Embley, T. M. (2003) Fungal hydrogenosomes contain mitochondrial heat-shock proteins. Mol Biol Evol 20, 1051-1061.

van der Giezen, M., Slotboom, D. J., Horner, D. S., Dyal, P. L., Harding, M., Xue, G. P., Embley, T. M. and Kunji, E. R. (2002) Con-

served properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. Embo J 21, 572-579.

Voncken, F., Boxma, B., Tjaden, J., Akhmanova, A., Huynen, M., Verbeek, F., Tielens, A. G., Haferkamp, I., Neuhaus, H. E., Vogels, G., Veenhuis, M. and Hackstein, J. H. (2002) Multiple origins of hydrogenosomes: functional and phylogenetic evidence from the ADP/ATP carrier of the anaerobic chytrid *Neocallimastix* sp. Mol Microbiol 44, 1441-1454.

Ward, B. L., Anderson, R. S. and Bendich, A. J. (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). Cell 25, 793-803.

Wernegreen, J. J. (2002) Genome evolution in bacterial endosymbionts of insects. Nat Rev Genet 3, 850-861.

Wiemann, M. and Görtz, H. D. (1991) Identification and localization of major stage-specific polypeptides of infectious *Holospora obtusa* with monoclonal antibodies. J Bacteriol 173, 4842-4850.

Williams, B. A. and Keeling, P. J. (2003) Cryptic organelles in parasitic protists and fungi. Adv Parasitol 54, 9-68.

Zhaxybayeva, O., Lapierre, P. and Gogarten, J. P. (2004) Genome mosaicism and organismal lineages. Trends Genet 20, 254-260.