Mitochondrial Evolution: Should I stay or should I go?

Jonathan Choy, Kenneth Liu, Catherine Tucker, Mitra Esfandirei, Steven Quayle

Canadian Institute for Advanced Research, Program in Evolutionary Biology, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada

The distinction between prokaryotic and eukaryotic cell structure is still accepted today as the most fundamental discontinuity in the living world. In the process of analyzing the newly sequenced bacterium *Rickettsia prowazekii* it was discovered through a BLAST search that a non-coding region of DNA showed high homology to the importin-α gene of eukaryotes. The genomes of *Rickettsia canada* and *Rickettsia rickettsii* were both found to contain a sequence homologous to importin-α as well. This sequence was found to have high homology when compared to the primitive protists *Nosema locustae* and *Reclinomonas americana,* which close ancestors to the lineage. The predicted protein sequence of *R. prowazekii*, *R. canada* and *R. rickettsii* contained the highly conserved amino acid motif cys-arg-glu-ala-thr-glu-…-ser-glu-val-glu-asn-asp-ala-tyr-ser. We believe this motif may be specific to a new lineage of prokaryotes, which we have termed the Eukobacteria. Through the data collected we propose a new model of mitochondrial evolution wherein one or more mitochondria escaped from their eukaryotic hosts and developed into the Eukobacteria.

Introduction

Resolving the order of events that occurred during the transition from prokaryotic to eukaryotic cells remains one of the greatest problems in cell evolution. For the past three decades, however, ideas about how these basic cell types are evolutionary related have changed drastically (Roger 1999). It is now generally accepted that most eukaryotes are fundamentally chimaeric, since different components of the eukaryotic cell have demonstrably different histories (Brown and Doolitle 1997). One view proposes that the endosymbiotic origin of mitochondria occurred relatively late in eukaryotic evolution and that several mitochondrion-lacking protist groups diverged before the establishment of the organelle (Bui et al. 1996). As well most evidence suggests that the mitochondrial endosymbiosis event took place prior to the divergence of all extant eukaryotes (Roger 1999).

Comparisons of small subunit ribosomal RNA (SSU rRNA) genes encoded on eukaryotic organellar and prokaryotic genomes indicate that mitochondria are specifically related to the *Rickettsia* subgroup of α-proteobacteria (Gray and Spencer). The *Rickettsia* are α-proteobacteria that only multiply in eukaryotic cells. *R. prowazekii* is the agent of epidemic, louse-borne typhus in humans (Gross 1996). That modern *Rickettsia* favour an intra-cellular lifestyle identifies these bacteria as the sort of organ-isms that might have initiated the endosymbiotic scenario leading to modern mitochondria (Margulis 1970).

Andersson et al. (1998) described the complete genome sequence of *Rickettsia prowazekii* as having 1,111,523 base pairs. Its gene content, like that of other parasitic eubacteria, has been reduced and tailored to suit its dependent lifestyle. Andersson et al. (1998) found that the *R. prowazekii* genome encodes 834 complete open reading frames, or DNA sequences that potentially specify protein sequences. Surprisingly, the *R. prowazekii* genome also contains the highest fraction of non-coding DNA (24%) found in any microbial genome so far, much of which may represent inactive genes that have been degraded by mutation, but have not yet been eliminated from the genome (Carlin 1999).

The DNAs of *Rickettsia* and mitochondria have highly derived genomes and are the products of several modes of reductive evolution. Both lack genes for metabolizing sugars in the absence of oxygen (anaerobic glycolysis), as well as all or most of the genes involved in synthesizing amino acids and nucleotides (Yang et al. 1997). The functional profiles of *Rickettsia* and mitochondria are strikingly similar, with production of ATP occurring in basically the same way in the two systems (Olsen and Woese 1996). *R. prowazekii* is more closely related to mitochondria than is any other bacterium whose genome has been investigated at this level of detail.

The current view holds that *Rickettsia* and mitochondrial genomes independently descended from an α-proteobacteria-like ancestor, each undergoing a separate process of reductive evolution (Lang et al. 1998). Evidence for this theory was largely based on the comparison of the organization of the same genes in the *R. prowazekii, E.coli* and *Reclimonas* mitochondrial genomes (Woese 1998). However, through analysis of non-coding regions in organisms from the *Rickettsia* genus we have gathered evidence that contradicts this theory and we have proposed a new theory in which some prokaryotes (*Rickettsia*) may have served as a primitive mitochondriate-like organism before separating from the eukaryotic lineage. It will be important to identify and explore the genomes of those minimally diverged, free-living α-proteobacteria that are specific but more distant relatives of both the *Rickettsiae* and mitochondria. Such genomes should yield additional clues relevant to the origin and evolution of mitochondria, a process that is central to the emergence of eukaryotic life.

Key words: mitochondria evolution, *Rickettsia prowazekii*, endosymbosis.

Address for correspondence and reprints: Steven Quayle, Canadian Institute for Advanced Research, Program in Evolutionary Biology, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, B3H 4H7, Canada.
Email: Steven-Q@is.dal.ca

Mol. Biol. Evol. 18(3):188-192. 2001

Materials and Methods

Sources of Sequences and DNA

The genomic sequence of *R. prowazekii* from 886-916kb was downloaded from the website for the National Centre for Biotechnology Information (accession number AJ235269).

DNA from *R. prowazekii* was a gift from Jaclyn Cleary (Universiteit van Amsterdam). DNA from *R. canada* and *R. rickettsii* were gifts from Michael Smith (University of Calgary). DNA from *Nosema locustae* and *Reclinomonas americana* were gifts from Martha Nelson (University of Hawaii).

Sequence Analysis

Searches for matches of nucleotide sequences of *R. prowazekii* in the nonredundant GenBank database were done using BLASTN (Altschul et al. 1997).

Searches for ORFs and amino acid sequence predictions were performed using the ExPASy Translate tool (http://www.expasy.ch/tools/dna.html). Hidden Markov model (HMM) searches of peptide sequences were performed at the website of the Protein families database (Pfam), version 6.1 (http://www.sanger.ac.uk/Software/Pfam/). Protein sequences were aligned using CLUSTAL W (Thompson, Higgins, and Gibson 1994) and shaded with Boxshade, version 3.21 (http://www.ch.embnet.org).

PCR Amplification

Degenerate primers were designed from the predicted protein products of two nearby ORFs in the pseudogene-rich region of the *R. prowazekii* genome between 886kb and 916kb. All PCR amplifications were performed using AccuTaq™ LA DNA Polymerase (Sigma) and the amplification products were analyzed by electrophoresis on agarose gel after staining with ethidium bromide. The optimal cycling profile varied between organisms, but the annealing step was carried out at approximately 53^{0}C and extension occurred for approximately 1.5 minutes at 72^{0}C.

Cloning and Sequencing

All PCR amplification products were cloned in electrocompetent *Escherichia coli* using the Zero Blunt™ TOPO® PCR Cloning Kit (Invitrogen). Clones carrying inserts were identified with blue/white selection. Sequencing was performed for at least three clones per PCR amplification product. Sequencing templates were prepared from overnight cultures of positive clones using standard protocols. The sequencing reaction was performed according to the protocol of ABI Prism Big Dye Terminator Sequencing Kit (Perkin Elmer). Extension products were purified by Sephadex G-50 (DNA grade, Pharmacia) and analyzed by the Genome Sequencing Center at Dalhousie University. All sequences obtained in this experiment have been deposited in the GenBank database with accession numbers AJ293329-AJ293333.

Results

Non-coding ORFs in *R. prowazekii*

ORFs from a 30 kilobases (kb) region located at position 886-916kb of *R. prowazekii* was chosen for this study because of its high in non-coding DNA and also possesses significantly higher G+C content than other non-coding areas of the genome. Four short ORFs ranging from 45 to 86 base pairs clustered in the 891-893kb region of *R. prowazekii* showed significant homology to the importin-α gene previously found only in eukaryotes.

IBB domain in *R. prowazekii*

Each of the four ORFs was conceptually translated using ExPASy, and these were then concatenated according to their respective positions in the genome. This predicted peptide sequence was searched in the Pfam database and the four combined sequences were assigned to the importin-β binding (IBB) domain family (accession number PF01749).

Interestingly, these four predicted amino acid sequences account for nearly 30% of the IBB domain in a non-overlapping fashion. Furthermore, due to the translation and database search methods used, the relative position of the four predicted sequences corresponds to the respective position of the ORFs within the genome. However, a 28 amino acid deletion in the proposed IBB domain was observed (fig. 1).

IBB in other *Rickettsia* and protists

Since *R. prowazekii* is an obligate intracellular bacterium, it

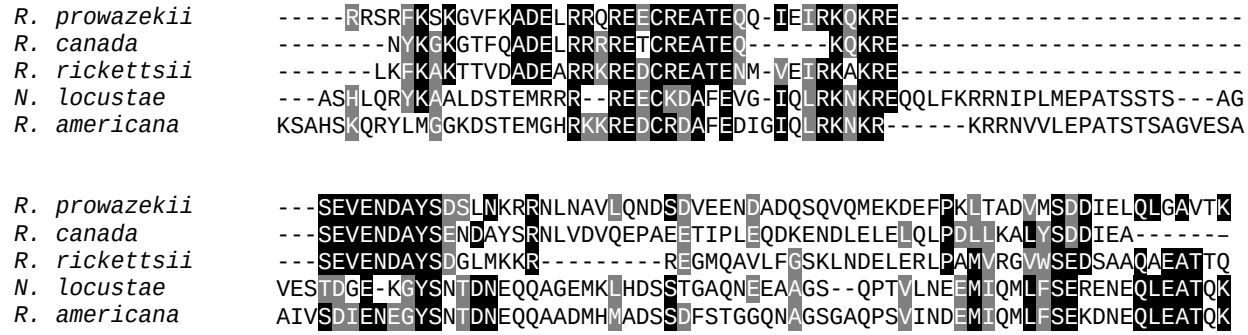

FIG. 1.-Alignment of the predicted peptide sequences of *Reclinomonas americana*, *Nosema locustae*, and three species of *Rickettsia*: *R. rickettsii*, *R. canada*, and *R. prowazekii*. These sequences were predicted from PCR-positive clones and were aligned with CLUSTAL W. Black boxes show residues that are identical between at least three sequences, while conservative substitutions are shaded grey.

is possible that it could have acquired the IBB domain from an eukaryote during a "recent" infection. To rule out this possibility, we designed degenerate PCR primers based on the predicted IBB protein product from *R. prowazekii* and performed a PCR screen in other *Rickettsia* species to look for the presence of this sequence. As expected, a fragment of approximately 500bp was observed in *R. prowazekii, R. canada* and *R. rickettsii*. (fig. 2)

The same pair of primers also amplified a product of approximately 450bp from two primitive protists, *N. locustae* and *R. Americana*, which were chosen due to their relative proximity to *Rickettsia* on the evolutionary tree (Roger 1999, Sicheritz et al. 1998). As a control, we used the same primers to look for an amplification product in *E. coli*, but no significant bands were found (fig. 2).

We then cloned the amplified PCR products in *E. coli*, and subsequently sequenced the insert. The DNA sequences were analyzed and amino acid sequences were again predicted using the ExPASy translate tool.

Alignment of these sequences revealed that there has been conservation of this IBB domain within the *Rickettsia* (fig. 1). In contrast, the homology of these sequences with *N. locustae* and *R. Americana* was less prominent. The 28 amino acid deletion was observed in all three *Rickettsia*, but was not present in the two protists sequenced here (fig. 1).

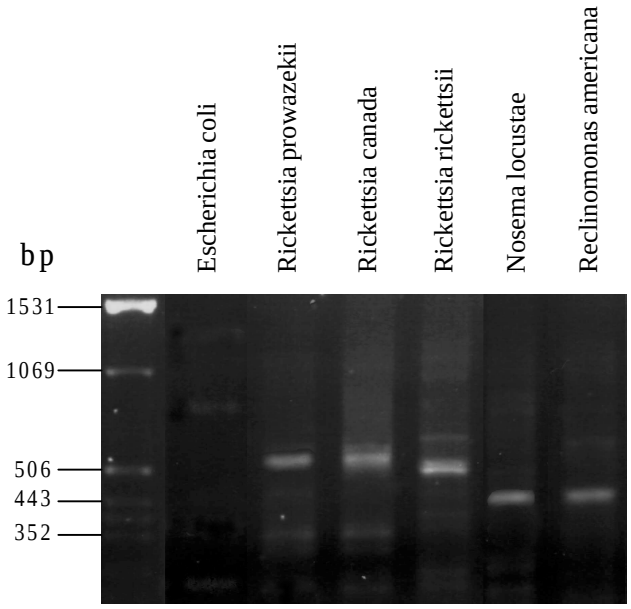

FIG.2-Agarose gel showing PCR products amplified from genomic DNA using degenerate primers designed from the predicted amino acid sequence of the *R. prowazekii* ORF's.

Discussion

The newly sequenced genome of *R. prowazekii* prompted our laboratory to explore certain sequences of unknown homology and function in the genome of this pathogen. A BLAST search on four small, non-coding regions of the genome showed significant homology to the importin-α gene, previously found only in eukaryotes. In order to determine if other members of the *Rickettsia* genus possessed the same importin-α sequence, we performed a PCR screen using degenerate primers designed from the sequence of the proposed IBB domain of importin-α found in *R. prowazekii*. Surprisingly, the genomes of *R. canada* and *R. rickettsii* both contained fragments that were amplified by this primer pair, suggesting the possibility that both of these organisms also carried IBB-like domains (fig. 2). These same primers were also used to amplify a slightly smaller fragment from *N. locustae* and *R. americana*, two primitive eukaryotes. *E. coli* was used as a control for the PCR screen to ensure that the level of degeneracy of the primer did not cause them to amplify unrelated regions of DNA (fig. 2).

As a further confirmation we cloned and sequenced the most prominant fragments amplified for each organism. We determined the predicted amino acid sequence of the ORF's in these fragments, and we aligned these sequences with that of the IBB-like domain of *R. prowazekii* (fig. 1). Significant homology between the *Rickettsia spp.* and eukaryotic sequences suggests that they might share a common ancestor.

Importin-α is responsible for recognizing the nuclear localization sequence of proteins that are targeted to the nucleus, and, in the presence of importin-β, shuttling these proteins into the nucleus (Gorlich and Mattaj, 1996). Bacteria do not have a nuclear compartment, thus it was surprising to discover a sequence in the non-coding region of *Rickettsia* that was homologous to the importin-β binding domain (IBB) of importin-α.

Further analysis of the predicted protein sequence of this region revealed homology to the protein sequence of IBB of importin-α of *N. locustae*. Further, the predicted protein sequence of *Rickettsia* importin-α contained a region predicted to be encoded by the amino acids cys-arg-glu-ala-thr-glu….ser-glu-val-glu-asn-asp-ala-tyr-ser. Although there is as yet no known function for this motif, and since it occurred only in the predicted protein sequence of *R. prowazekii, R. canada* and *R. rickettsii* importin-α, we propose that it may be specific to a new lineage of prokaryotes, the Eukobacteria, and may aid in the ancestral determination of bacteria. We have named this domain the genesis motif.

Since *R. prowazekii* is an intracellular pathogen and may acquire host genes through gene transfer, it is possible that the *Rickettsia* bacteria studied in this report acquired the importin-α gene through this mechanism. However, we believe that the importin-α gene was instead transferred to a primitive mitochondria which later diverged to form the *Rickettsia* genus. We have identified remnants of importin-α IBB in 3 species of Rickettsia. The homology of these sequences and their predicted protein sequences is consistent with this gene having been acquired before the divergence of *R. prowezekii*, commonly believed to be the ancestor of mitochondria. Therefore, the most parsimonious explanation for this observation is that a common ancestor acquired this gene from a nucleated organism around the time we believe the exclusion event occurred.

It is believed that the genesis of the eukaryotic lineage arose from the endosymbiosis of a primitive, mitochondrial-like bacteria into a nucleated cell (Gray, 1993). Our data presented here suggest that a subset of bacteria may have separated from the eukaryotic lineage after the endosymbiotic event. Under this paradigm, changes in selection pressures would have favored the release of a primitive mitochondria from the cytosol of eukaryotes, giving rise to a new group, the Eukobacteria.

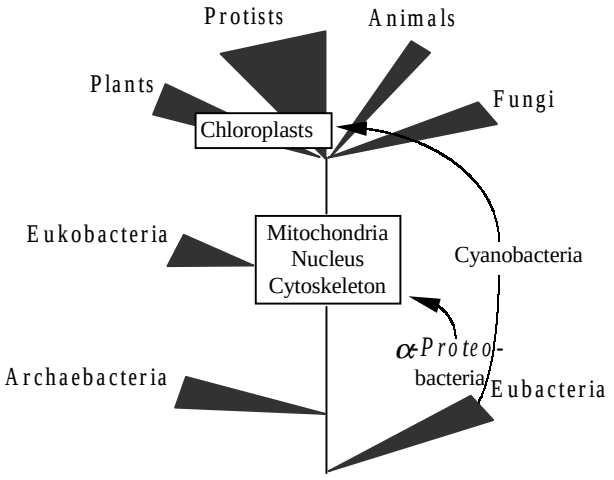

FIG.3-An updated view of early events in evolution. The presence of nuclear genes in *Rickettsia* (fig. 2) suggests that mitochondria were formed by members of the α-proteobacteria, some of which later escaped from their eukaryotic host to establish the Eukobacteria.

Furthermore, *R. prowazekii* is believed to be the most common ancestor to this mitochondria-derived, primitive bacteria.

An outline of our model, termed the Mitochondrial Exclusion theory, is presented in Figure 3. The endosymbiotic event would have initially favored the association of a mitochondrial-like bacteria with a nucleated organism. However, changes in selection pressure may have eventually led to the dissociation of the mitochondria from the organism. This is supported by the existence of amitochondriate protists, such as *N. locustae*, that possess remnants of mitochondrial genes, suggesting that they once harbored a primitive version of this organelle that was later lost. If dissociation of the mitochondria was associated with a limited dependence of this primitive organelle on the host, nucleated organism, the persistence of a mitochondrial-derived organism may have given rise to a new kingdom of organisms, which we termed the Eukobacteria.

Further support for this theory comes from the ATP-ADP transporter required by aerobic organisms to import ATP from the mitochondria in exchange for ADP (Gray, 1998). Although coded in the nuclear genome of many eukaryotes, in bacteria this protein has been observed in only *Rickettsia* and *Chlamydia*. It is believed that since *Rickettsia* and *Chlamydia* are both intracellular pathogens, acquistion of the ATP-ADP transporter from the host DNA may have been beneficial to "hijack" host ATP in order to maintain metabolism of these pathogens while they were associated with the host (Kurlans, 1992). We believe the presence of the ATP-ADP transporter in these bacteria strengthens our Mitochondrial Exclusion theory since it demonstrates that gene transfer from the nucleus to the mitochondria does indeed occur. Moreover, since gene transfer from the host nucleus to an intracellular pathogen is a rare event, we believe that transfer from the nucleus to the mitochondria may be a more probable event since the two entities would have been associated for a greater amount of time than a pathogen and host. In this study we did not search the *Chlamydia* genome for the genesis motif, however, future analysis may prove intriguing.

Through analysis of non-coding regions in organisms from the Rickettsia genus and comparison of these sequences with those of two protists, we have found evidence suggesting that some prokaryotes separated from the eukaryotic lineage after the initial endosymbiotic event that gave rise to mitochondria. These prokaryotes may have originally served as primitive mitochondria before separating from their host organism. Therefore, we have assigned these prokaryotes to a new kingdom, Eukobacteria. The identification of other prokaryotes belonging to this kingdom will depend on genetic analysis of further prokaryotes and will be expedited by the current advances in genetic and molecular biology.

Acknowledgments

Many thanks to C. Kurland and H. Winkler for helpful comments on earlier versions of the manuscript. This work was supported by grants of Canadian Bacterial Disease Network, Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. and LIPMAN, D. J. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**:3389-402.

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**:133-40.

BROWN, J. R. and DOOLITTLE, W. F. 1997 Archaea and the prokaryote-to-eukaryote transition. Microbiol Mol Biol Rev **61**:456-502.

BUI, E. T., BRADLEY, P. J. and JOHNSON, P. J. 1996 A common evolutionary origin for mitochondria and hydrogenosomes. Proc Natl Acad Sci U S A **93**:9651-6.

EMBLEY, T. M. and MARTIN, W. 1998 A hydrogen-producing mitochondrion. Nature **396**:517-9.

GORLICH, D. and MATTAJ, I. W. 1996 Nucleocytoplasmic transport. Science **271**:1513-8.

GRAY, M. W. 1993 Origin and evolution of organelle genomes. Curr Opin Genet Dev **3**:884-90.

GRAY, M. W. and Spencer D.F. 1996 Organeller evolution. pages 109-126 in D.M. Roberts and M.A. Collins, eds. Evolution of microbial life: 54th symposium of the society for geneeral microbiology. Cambridge University Press, Cambridge.

GRAY, M. W. 1998 Rickettsia, typhus and the mitochondrial connection. Nature **396**:109-10.

GRAY, M. W. 2000 Mitochondrial genes on the move. Nature **408**:302-3, 305.

GROSS, L. 1996 How Charles Nicolle of the Pasteur Institute discovered that epidemic typhus is transmitted by lice: reminiscences from my years at the Pasteur Institute in Paris. Proc Natl Acad Sci U S A **93**:10539-40.

KARLIN, S., BROCCHIERI, L., MRAZEK, J., CAMPBELL, A. M. and SPORMANN, A. M. 1999 A chimeric prokaryotic ancestry

of mitochondria and primitive eukaryotes. Proc Natl Acad Sci U S A **96**:9190-5.

KURLAND, C. G. 1992 Evolution of mitochondrial genomes and the genetic code. Bioessays **14**:709-14.

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**:493-7.

MARGULIS, L. 1970 Origin of Eukaryotic Cells. Yale University Press, New Haven.

OLSEN, G. J., WOESE, C. R. and OVERBEEK, R. 1994 The winds of (evolutionary) change: breathing new life into microbiology. J Bacteriol **176**:1-6.

ROGER, A.J. 1994 Reconstructing early events in eukaryotic evolution. American Naturalist **154**: suppl. 146-163.

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**:545-51.

THOMPSON, J. D., HIGGINS, D. G. and GIBSON, T. J. 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673-80.

WEISBURG, W. G., WOESE, C. R., DOBSON, M. E. and WEISS, E. 1985 A common origin of rickettsiae and certain plant pathogens. Science **230**:556-8.

WOESE, C. R. 1987 Bacterial evolution. Microbiol Rev **51**:221-71.

YANG, D., OYAIZU, Y., OYAIZU, H., OLSEN, G. J. and WOESE, C. R. 1985 Mitochondrial origins. Proc Natl Acad Sci U S A **82**:4443-7.