Classification: Major: Biological Sciences; Minor: Evolution

The Mitochondrial Genome of *Paraspadella gotoi* is Highly Reduced and Reveals that Chaetognaths are a Sister-Group to Protostomes

Kevin G. Helfenbein[1,2*] , H. Matthew Fourcade[2], Rohit G. Vanjani[1], and Jeffrey L. Boore[1,2,3]

[1] Department of Biology, University of Michigan, 830 N. University Ave., Ann Arbor, MI 48109 USA

[2] DOE Joint Genome Institute and Lawrence Berkeley National Laboratory, 2800 Mitchell Drive, Walnut Creek, CA 94598 USA

[3] Department of Integrative Biology, University of California Berkeley, Berkeley, CA 94720

[*] Present address: American Museum of Natural History, Department of Invertebrate Zoology, Central Park West at 79th St., New York, NY 10024

Corresponding author: Jeffrey L. Boore, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek CA 94598, telephone 925-296-5691, Fax: 925-296-5666, JLBoore@lbl.gov

Keywords: Chaetognath, Paraspadella, phylogeny, mitochondria, evolution, gene loss

Manuscript information: 22 text pages including the title page, abstract, references, figure legends and 4 figures (no tables). All text characters are 26,643 plus Fig. 1 = 1,380, Fig. 2 = 1,596, Fig. 3 = 3,480, Fig. 4 = 2,940. Total length of the manuscript is estimated at 36,039 characters.
GenBank accession number AY619710

ABSTRACT

We report the first complete mitochondrial (mt) DNA sequence from a member of the phylum Chaetognatha (arrow worms). The *Paraspadella gotoi* mtDNA is highly unusual, missing 23 of the genes commonly found in animal mtDNAs, including *atp6*, which has otherwise been found universally to be present. Its 14 genes are unusually arranged into two groups, one on each strand. One group is punctuated by numerous non-coding intergenic nucleotides, while the other group is tightly packed, having no non-coding nucleotides, leading to speculation that there are two transcription units with differing modes of expression. The phylogenetic position of the Chaetognatha within the Metazoa has long been uncertain, with conflicting or equivocal results from various morphological analyses and rRNA sequence comparisons. Comparisons here of amino acid sequences from mitochondrially encoded proteins gives a single most parsimonious tree that supports a position of Chaetognatha as sister to the protostomes studied here. From this, one can more clearly interpret the patterns of evolution of various developmental features, especially regarding the embryological fate of the blastopore.

INTRODUCTION

Chaetognaths comprise a phylum of marine invertebrates, ranging in size from 0.5-12 cm, that prey upon small animals such as copepods and immature fish. These animals have prominent lateral fins and a tail fin from whence their common name "arrow worm" is derived. The phylogenetic position of this phylum within the Metazoa has long been controversial. Comparisons of morphological features have yielded very different results, sometimes placing the Chaetognatha as deuterostomes (e.g., sea urchins, vertebrates) (1), but otherwise as protostomes (e.g., flies, clams) (2). Nuclear ribosomal RNA (rRNA) sequence comparisons have consistently suggested that chaetognaths are not deuterostomes; however, their sequences are highly divergent from those of other metazoans and highly skewed in base composition, causing some to doubt the reliability of these analyses, and leaving the phylogenetic position of Chaetognatha equivocal (3).

Comparisons of complete mtDNA sequences have frequently been useful for reconstructing ancient phylogenetic relationships and serve as a model for the processes of genome evolution. From the nearly universal gene content of sequenced animal mtDNAs (4, 5), it is simple to deduce that the last common ancestor of triploblast animals had a mtDNA encoding 37 genes: 13 protein encoding genes, two rRNA genes, and 22 tRNA genes. Although this gene content varies little among animals (4), the arrangement of these genes varies for many groups. For some mtDNAs, genes are encoded in only one strand, whereas in others they are divided between the two strands. For the few cases where it has been studied, transcription of these genes proceeds uninterrupted around the circular DNA to make a single RNA that is then cut into individual gene specific fragments, in at least some cases by excision of the tRNAs from the transcript (6).

However, it is common for mtDNAs to have protein encoding genes that abut without an intervening tRNA gene, and it remains unclear how (or if) their mRNAs are separated after transcription.

Complete sequences have been determined for approximately 230 animal mtDNAs, but the taxonomic sampling has been highly biased, with about 70% being from vertebrates alone. For many whole phyla, there have been no studies of mtDNA sequences, or of gene content, arrangement, or expression. This is the first report of a complete mtDNA sequence from any member of the phylum Chaetognatha, revealing several features that are novel among animal mtDNAs. The large scale sequence comparisons enabled by this data strongly support a position of Chaetognatha as sister-group to the radiation of protostome animals, at least as sampled here with annelids, mollusks, phoronids, brachiopods, and arthropods.

METHODS AND MATERIALS

DNA Amplification, Sequencing, and Informatics. A preparation *P. gotoi* total DNA was a gift of Nori Satoh, Kyoto University. Conserved regions within the mitochondrial genes *cox1* and *cox3* were amplified by PCR using primers designed to match conserved regions. These products were sequenced and specific primers were designed facing "out" from each gene and used in long PCR with the enzyme LA-PCR from Takara, resulting in amplifications between a forward *cox1* primer and reverse *cox3* primer and vice versa (Fig. 1). The long PCR products were sequenced by a combination of primer walking using an ABI 377 sequencer and random shotgun sequencing of sheared PCR products.

The final assembly of all reads was performed with Sequencher (GeneCodes). *P. gotoi* protein and rRNA encoding genes were annotated based on comparisons with homologous genes from other animal species. Searches for tRNA genes were performed using tRNA scan-SE (7) or, in cases where this failed, were found by eye with the aid of GCG (8).

Sequences. For phylogenetic analyses, we chose a broad representation of taxa from the available complete mtDNA sequences. We omitted those taxa with the most highly divergent sequences, an approach shown previously to be highly effective for reconstructing metazoan phylogeny using the small nuclear ribosomal RNA (9). The following sequences were retrieved from GenBank: *Asterina pectinifera* (Echinodermata) (accession number NC_001627), *Balanoglossus carnosus* (Hemichordata) (NC_001887), *Daphnia pulex* (Arthropoda) (NC_000844), *Drosophila yakuba* (Arthropoda) (NC_001322), *Florometra serratissima* (Echinodermata) (NC_001878), *Homo sapiens* (Chordata) (NC_001807), *Katharina tunicata* (NC_001636) (Mollusca), *Laqueus rubellus* (Brachiopoda) (NC_002322), *Limulus polyphemus* (Arthropoda) (NC_003057), *Locusta migratoria* (Arthropoda) (NC_001712), *Loligo bleekeri* (Mollusca) (NC_002507), *Lumbricus terrestris* (Annelida) (NC_001673), *Metridium senile* (Cnidaria) (NC_000933), *Mustelus manazo* (Chordata) (NC_000890), *Paracentrotus lividus* (Echinodermata) (NC_001572), *Phoronis architecta* (Phoronida) (AY368231), *Pichia canadensis* (Fungi) (NC_001762), *Platynereis dumerilii* (Annelida) (NC_000931), *Podospora anserina* (Fungi) (NC_001329), *Terebratalia transversa* (Brachiopoda) (NC_003086), *Terebratulina retusa* (Brachiopoda) (NC_000941).

Sequence Alignments and Phylogenetic Analysis. Inferred amino acid sequences of the 11 protein encoding genes held in common among the 19 ingroup and three outgroup taxa were aligned using ClustalW. Eight of these genes (*cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad3*, *nad4*, and *nad5*) were subjectively judged to be of reliable and unambiguous alignment. Pairwise alignments were done in slow mode with an open gap penalty of 10 and an extend gap penalty of 1 using the BLOSUM similarity matrix. The multiple alignment used the same parameters as the pairwise plus a "delay divergent sequences" setting of 40%. Ambiguously aligned positions at the termini of each gene alignment were trimmed.

The phylogenetic tree was obtained by parsimony analysis using a heuristic search with 10,000 random additions in PAUP* version 4.01b10 (10). The bootstrap values were calculated using 500 bootstrap replicates, each with 10 random addition sequence replicates. Bremer support values were calculated using Tree Rot (11), and the statistical significance of this value found for the protostomes + chaetognath clade was tested using the Templeton (Wilcoxon's signed ranks) test (12).

RESULTS AND DISCUSSION

Genome Evolution. The *Paraspadella gotoi* mtDNA is remarkably small, consisting of only 11,423 bp. It contains only 14 genes rather than the 37 almost universally found in the mitochondrial genomes of other triploblastic animals. Missing are 21 of the 22 tRNA genes normally present and the protein encoding genes *atp6* and *atp8*. The only tRNA gene that can be identified by searching for sequences with tRNA-like potential

secondary structures is *trnM*, which is also the case for the mtDNAs of the diploblasts *Sarcophyton glaucum* and *Renilla kolikeri*, and nearly the same as for *Metridium senile* mtDNA, which has genes for only two tRNAs, *trnM* and *trnW* (13-15). These three animals are in the phylum Cnidaria, wherein this gene loss must have occurred independently to their loss in the Chaetognatha. One might expect that the methionyl-tRNA would be the least "expendable" due to its unique role in initiating mitochondrial proteins with formyl-methionine (16).

The long-PCR reactions that amplified the mtDNA in two overlapping fragments (Fig. 1) produced strong, singular bands during electrophoretic analysis, bolstering confidence that this completely represents this circular molecule. However, it is not possible to rule out the alternative that the mtDNA contains a widely separated duplication of either *cox1* or *cox3* such that sequencing from the 3' end of one duplicate gene copy to the 5' end of the other would yield an apparently complete but truncated version of the mtDNA. The genes appearing to be absent, then, would actually be located between the two duplicated copies in an unstudied portion of the mtDNA. If that were the case, it would compel us to the remarkable coincidence that all of the 21 unfound tRNAs and one of the two unfound protein encoding genes (*atp8*) comprise the exact set of genes previously found to be missing in any animal mtDNA.

Genes encoding tRNAs capable of carrying the 19 amino acids other than methionine must be present somewhere in the *P. gotoi* cell. Since they do not seem to be encoded in this DNA molecule, there remain two possible alternatives. First, the genes could be encoded by a second DNA in the mitochondrion. Cases of second, or multiple, mtDNAs include linear plasmids in fungi (17), and diverse mtDNAs in the nematode *Globdera*

pallida, each with one or a few genes and much non-coding sequence (18). Neither the fungal plasmids nor any of the sequenced *G. pallida* mtDNAs contain tRNA genes. Secondly, the genes may be in the nucleus, with their products imported from the cytosol into the mitochondria. There are examples of tRNA import into the mitochondrion from diverse taxonomic groups. Apicomplexans and trypanosomatids, which encode no tRNAs in their mtDNAs (19), and plants, encoding only some tRNAs in their mtDNAs, import multiple tRNAs from the cytosol into the mitochondrion (20). In yeast, tRNA(K) is imported from the cytosol redundantly (i.e., there is a functional, mitochondrially encoded copy) (21), and in marsupials the same tRNA is also imported, although here the mitochondrial copy is a pseudogene (22).

The *atp8* gene encodes subunit 8 of the F_o portion of the enzyme ATP synthase and, in yeast, this has been implicated in structure and assembly of the ATPase and for coupling proton transport from the F_o sector to the F_1 sector where ATP synthesis occurs (reviewed in 23). In metazoans, Atp8 is the smallest mitochondrially encoded protein, only about 50-65 amino acids long, and only about a half dozen of these amino acid residues are well conserved across animal mtDNAs. Among the major animal taxa that have been sampled, *atp8* has been lost from the mtDNA independently in bivalve mollusks (e.g., 24), secernentean nematodes (25), and platyhelminths (26). Future mtDNA sequences may more clearly define the limits of this loss within these taxa and, as in the case of this chaetognath, reveal more taxa with this condition. Our attempts to find a copy by BLAST search of *atp8* in the nuclear genome of *C. elegans* (one of the secernentean nematodes) have failed. It may be that the ATP synthase of some metazoans can function acceptably without this protein.

Atp6 is directly involved in ATP synthesis (27). Although *atp6* is encoded in all other sequenced animal mtDNAs, it is not encoded in some non-metazoan mtDNAs. This is the case in *Chlamydomonas reinhardtii,* where *atp6* is a nuclear encoded gene (28). A barrier to movement of protein-encoded genes from the mtDNA to the nucleus may be the variation in the genetic code, especially the use of TGA to specify tryptophan in mitochondria instead of as a stop codon as in the "universal" genetic code. It is notable that *C. reinhardtii's* mitochondrial mRNAs are translated with the standard code, wherein TGA encodes a stop signal, so this would not apply in that case. However, for *P. gotoi*, TGA encodes tryptophan; in fact, TGA is the preferred tryptophan codon, 65 to 10, over TGG in the chaetognath mitochondrial protein encoding genes. Thus, in order for *atp6* to have transferred to the nucleus, either (a) the gene would not have contained any TGA codons when in the mtDNA, or (b) the necessary A —> G mutations to change all TGAs to TGGs took place prior to the rise of a disabling mutation after transfer to the nucleus. As mentioned above for tRNAs, it is possible that the missing protein genes may be encoded by a second DNA in the mitochondrion rather than by the nuclear genome. Alternatively, in the case of either *atp6* or *atp8*, the genes may be lost; perhaps their function has become dispensable or been subsumed by other proteins.

The organization of the genes in the *P. gotoi* mtDNA is unique among metazoan mtDNAs. Unlike any other studied animal mtDNA, the genes encoded in each strand are grouped together, as shown in Fig. 2. The protein genes in the clockwise orientation are separated by an unusually large number of nucleotides (Figs. 2 and 3) for an animal mtDNA. The protein genes in the opposite transcriptional orientation (shown counterclockwise in Fig. 2) are not separated at all. In fact, although we have annotated

these genes such that they end on abbreviated stop codons (see below and Fig. 3), the reading frames of *nad6, cob,* and *nad4* are open such that they could otherwise overlap the downstream gene. It has been shown that overlapping *atp8/atp6* and *nad4L/nad4* gene pairs form dicistronic mRNAs for some taxa (29), perhaps due to the smaller genes' (*atp8* and *nad4L*) inability to bind solo to the ribosome (30). This would not be true for *cob*, which is not particularly small for a mitochondrially encoded protein. These differences suggest that the transcripts of the two strands may be processed differently at one or more steps prior to translation, or that translation itself is different for the two sets of genes.

It has been proposed that the tRNA secondary structures that form in the primary transcript act as endonucleolytic targets, with their excision liberating both the tRNAs themselves as well as the intervening messages (6). Furthermore, it has been proposed that in lieu of a tRNA, a tRNA-like structure in a primary transcript can serve this function (29). Such structures have been noted for some mtDNAs (31), but in many cases there are genes abutting without any obvious, intervening potential secondary structures, and that is the case with the *P. gotoi* mitochondrial genome. We have annotated some genes to end on a 'T' in the first codon position, inferring that these are modified into complete TAA stop codons by polyadenylation of the mRNA after transcript cleavage. This appears to be common in mitochondrial systems (6), and for *P. gotoi* there seems to be no reasonable alternative, since some genes would have to overlap extensively to reach the first in-frame stop codon. However, it is not obvious what might be acting to signal cleavage of the polycistron at these precise locations. The

large number of these cases in the *P. gotoi* mtDNA serves to highlight this issue as one of the major unsolved problems of mitochondrial molecular biology.

Phylogenetics. Our analysis recovered a single most parsimonious tree of 16,267 steps (Fig. 4) that robustly supports the Chaetognatha—represented here by *P. gotoi*—as the sister group to the protostomes included in this study. This corroborates the result of Nielsen and coworkers (2), although their analysis placed the phoronids and brachiopods as deuterostomes, an untenable position given the results of independent molecular data sets (32, 33). The shortest alternative tree that does not include a monophyletic grouping of protostomes and chaetognaths requires 32 extra steps and is rejected by the Templeton test (12) in favor of the shortest tree uniting the groups (N= 116, z=-2.9711, P=0.003). Little variation was noted in any of the branch lengths in this tree as reconstructed by parsimony criteria. There are some embryological features recently noted (34) that also support the placement of chaetognaths as or near protostomes rather than as deuterostomes. Chaetognaths have a tetrahedral 4-cell embryo and an arrangement of blastomeres at the 4-cell stage corresponding to future body axes, as is found in some protostomes (i.e., Mollusca, Annelida, and some crustaceans).

Thus, morphological features that have been inferred to place the Chaetognatha within the deuterostomes, e.g., secondary mouth formation, mesoderm derived directly from the archenteron (1), and having a trimeric coelom [although this has been recently reevaluated in chaetognaths (35)] have either evolved independently in these groups or are retained from a more ancient lineage (i.e., are sympleisiomorphies). If the latter is the case, then one might infer that having the mouth arise from the blastopore during

embryological development is a shared and derived condition (i.e., a synapomorphy) of the protostomes, with mouth formation via a secondary opening in the embryo being ancestral for triploblastic animals, and being retained by chaetognaths and deuterostomes, as has been suggested by Peterson and Eernisse (36).

In their study (36), chaetognaths were placed either as the sister taxon to the clade of molting animals referred to as the Ecdysozoa (9) (using morphological characters), or within the Ecdysozoa (using small nuclear ribosomal RNA data or a combination of these data). The Bremer support values for these hypotheses of relationship were low, however, with the small nuclear ribosomal RNA sequence results described as "unstable", possibly due to extreme GC bias. The chaetognath sequences reported here show no unusual base compositional bias among animal mtDNA sequences. We specifically tested their hypothesis of a Chaetognatha + Ecdysozoa clade against the most parsimonious tree found, which increased the tree length by 27 steps; thus, our data are inconsistent with this view. (Ecdysozoa is represented by only arthropods in this study.)

A recent *Hox* gene survey in the chaetognath *Spadella cephaloptera* (38) led the authors to speculate that chaetognaths diverged from the triploblast lineage prior to the deuterostome/protostome split. We would suggest however, that that study illustrates the difficulty of using paralogous genes to reconstruct the evolutionary history of species; indeed, there is no actual phylogenetic analysis included therein. That the few morphological characters that ally the chaetognaths to the protostomes (see 2) and the various morphological features that are unique to the chaetognaths open the door to these speculations is undeniable. However, only strongly supported phylogenetic analyses, rather than narrative *ad hoc* hypotheses, will enable us to understand the relationships of

the animal phyla. We look forward to continued tests of the phylogenetic position of the chaetognaths, and new tests of other poorly studied animal phyla, as more complete animal mitochondrial genome sequences become available.

ACKNOWLEDGEMENTS

Thanks to Nori Satoh of Kyoto University for the gift of *P. gotoi* DNA, and to J. R. Macey for help with the Templeton test. This work was supported by DEB-9807100 from NSF and partly performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, Lawrence Berkeley National Laboratory, under contract No. DE-AC03-76SF00098.

LITERATURE CITED

1. Brusca, R. C. & Brusca, G. J. (1990) *Invertebrates* (Sinauer, Sunderland, MA).

2. Nielsen, C., Scharff, N. & Eibye-Jacobsen, D. (1996) *Biol. J. Linn. Soc. Lond.* **57,** 385-410.

3. Giribet, G., Distel, D. L., Polz, M., Sterrer W. & Wheeler, W. C. (2000) *Syst. Biol.* **49,** 539-562.

4. Boore, J. L. (1999) *Nucl. Acids Res.* **27,** 1767-1780.

5. Boore, J. L. http://www.jgi.doe.gov/programs/comparative/MGA_Source_Guide.html

6. Ojala, D., Montoya, J. & Attardi, G. (1981) *Nature* **290,** 470-474.

7. Lowe, T. M. & Eddy, S. R. http://www.genetics.wustl.edu/eddy/tRNAscan-SE.

8. Devereux, J., Haberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res*. **12,** 387-395.

9. Aguinaldo, A. M., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A. & Lake, J. A. (1997) *Nature* **387,** 489–493.

10. Swofford, D. L. (1998) *Phylogenetic analysis using parsimony (*and other methods)* (Sinauer, Sunderland Massachusetts) Beta Version 4.0b1.

11. Sorenson, M. D. (1996) *TreeRot.* (University of Michigan, Ann Arbor).

12. Templeton A. R. (1983) *Evolution* **37,** 221-244.

13. Beagley, C. T., Macfarlane, J. L., Pont-Kingdon, G. A., Okimoto, R., Okada, N. & Wolstenholme, D. R. (1995) in *Progress in Cell Research* **5,** 149-153.

14. Beagley, C. T., Okimoto, R. & Wolstenholme, D. R. (1998) *Genetics* **148,** 1091-1108.

15. Beaton, M. J., Roger, A. J. & Cavalier-Smith, T. (1998) *J. Mol. Evol.* **47,** 697-708.

16. Smith, A. E., & Marker, K. A.(1968) *J. Mol. Biol.* **38,** 241-243.

17. Meinhardt, F., Kempken, F., Kämper & Esser, K. (1990) *Curr. Genet.* **17,** 89-95.

18. Armstrong, M. R., Blok, V. C. & Phillips, M. S. (2000) *Genetics* **154,** 181-192.

19. Tan, T. H., Pach, R., Crausaz, A., Ivens, A. & Schneider, A. (2002) *Mol. Cell. Biol.* **22,** 3707-3716.

20. Dietrich, A., Weil, J. H. & Maréchal-Drouard, L. (1992) *Annu. Rev. Cell Biol.* **8,** 115-131.

21. Martin, R. P., Schneller, J. M., Stahl, A. J. C. & Dirheimer, G. (1977) *Nucl. Acids Res.* **4,** 3497-3510.

22. Dörner, M., Altmann, M., Pääbo, S. & Mörl, M. (2001) *Mol. Biol. Cell* **12,** 2688-2698.

23. Devenish, R. J., Prescott, M., Roucou, X. & Nagley, P. (2000) *Biochim. Biophys. Acta* **1458,** 428-442.

24. Hoffmann, R. J., Boore, J. L. & Brown, W. M. (1992) *Genetics* **131,** 397-412.

25. Okimoto, R., Macfarlane, J. L., Clary, D. O. & Wolstenholme, D. R. (1992) *Genetics* **130,** 471-498.

26. Le, T. H., et al. (2000) *Mol. Biol. Evol.* **17,** 1123-1125.

27. Vik, S. B., Patterson, A. R. & Antonio, B. J. (1998) *J. Biol. Chem.* **273,** 16229-16234.

28. Funes, S., Davidson, E., Gonzalo Claros, M., van Lis, R., Pérez-Martínez, X., Vázquez-Acevedo, M., King, M. P. & González-Halphen, D. (2002) *J. Biol. Chem.* **277,** 6051-6058.

29. Amalric, F., Merkel, C., Gelfland, R. & Attardi, G. (1978) *J. Mol. Biol.* **118,** 1-25.

30. Tanmaan, J. -L. (1999) *Biochim. Biophys. Acta* **1410,** 103-123.

31. Boore, J. L. & Brown, W. M. (1994) *Genetics* **138,** 423-443.

32. Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M. A., Liva, S. M., Hillis, D. M. & Lake, J. A. (1995) *Science* **267,** 1641-1643.

33. Helfenbein, K. G. & Boore, J. L. (2004) *Mol. Biol. Evol.* **21,** 153-157.

34. Shimotori, T. & Goto, T. (2001) *Develop. Growth Differ.* **43,** 371-382.

35. Kapp, H. (2000) *Zool. Antz.* **239,** 263-266.

36. Peterson, K. J. & Eernisse, D. J. (2001) *Evol. Dev.* **3,** 170-205.

37. Maddison, W. P., and Maddison, D. R. 2000. *MacClade4: Analysis of Phylogeny and Character Evolution.* Sinauer Associates, Sunderland, Mass.

38. Papillon, D., Perez, Y., Fasano, L., Le Parco, Y. & Caubit, X. (2003) *Dev. Genes Evol.* **213,** 142-148.

FIGURE LEGENDS

Figure 1

Electrophoretic gels of PCR products of the *P. gotoi* mtDNA. (A) "1 kb plus" ladder followed by a PCR fragment of 6,578 bp generated by using the *cox1* reverse primer and *cox3* forward primer. (B) PCR fragment of 4,054 bp generated by using the *cox1* forward primer and *cox3* reverse primer followed by a mass standard, then a 1 kb standard. Exact sizes were determined by sequencing.

Figure 2

Gene map of the *Paraspadella gotoi* mtDNA. Genes are transcribed left to right as they are read—*rrnL* to *rrnS* clockwise, *nad6* to *nad5* counterclockwise, as marked by the arrows. Gene sizes are not to scale. Numerals outside of the map indicate the number of nucleotides intervening between each pair of genes; those flanking *rrnS* and *rrnL* are estimates based on the extent of alignment of these genes with those of other animals, but remain uncertain.

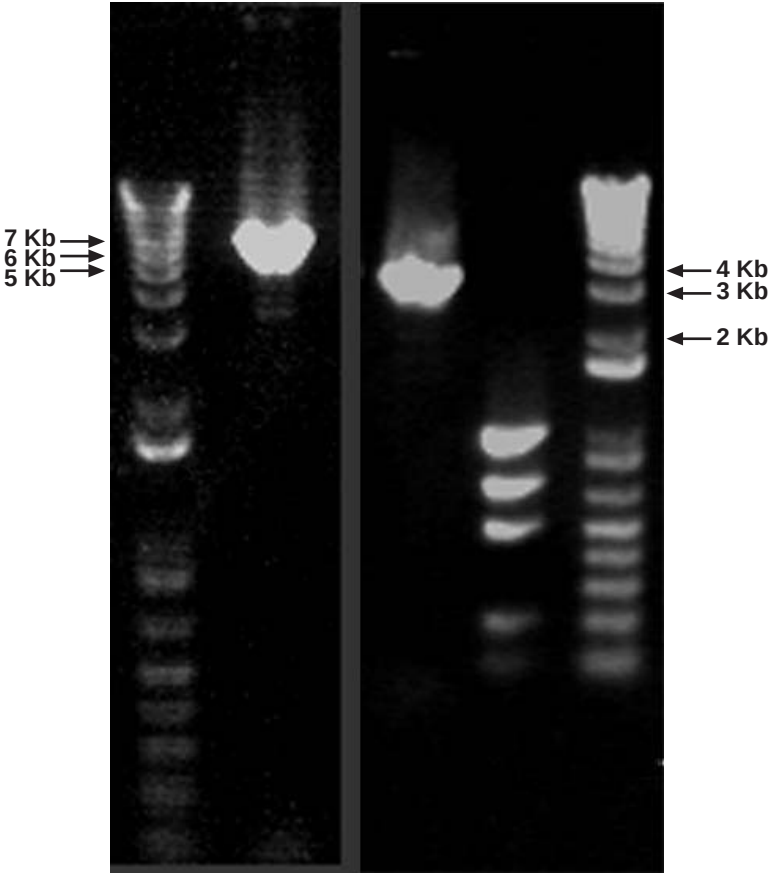
Figure 3

Abbreviated mtDNA sequence of *P. gotoi* shown graphically linearized at the beginning of *cox1*. Numerals within the slash marks indicate number of omitted nucleotides. Non-coding nucleotides, including the largest non-coding region, are underlined. A dart (>) marks the last nucleotide for each gene and indicates the direction of transcription. Nucleotides of termination codons of protein coding genes, complete or abbreviated (see text), are underscored with carats (^). All genes are inferred to initiate with formyl-

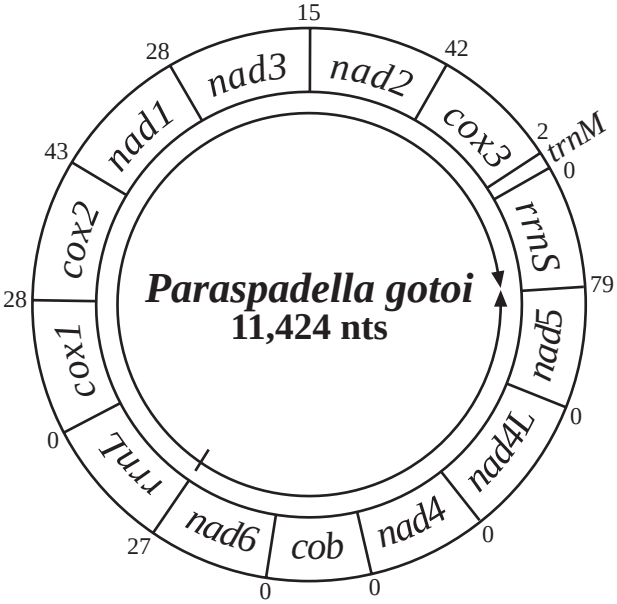
methionine regardless of the actual start codon; thus, this amino acid is shown in parentheses when not conforming to the genetic code.

Figure 4

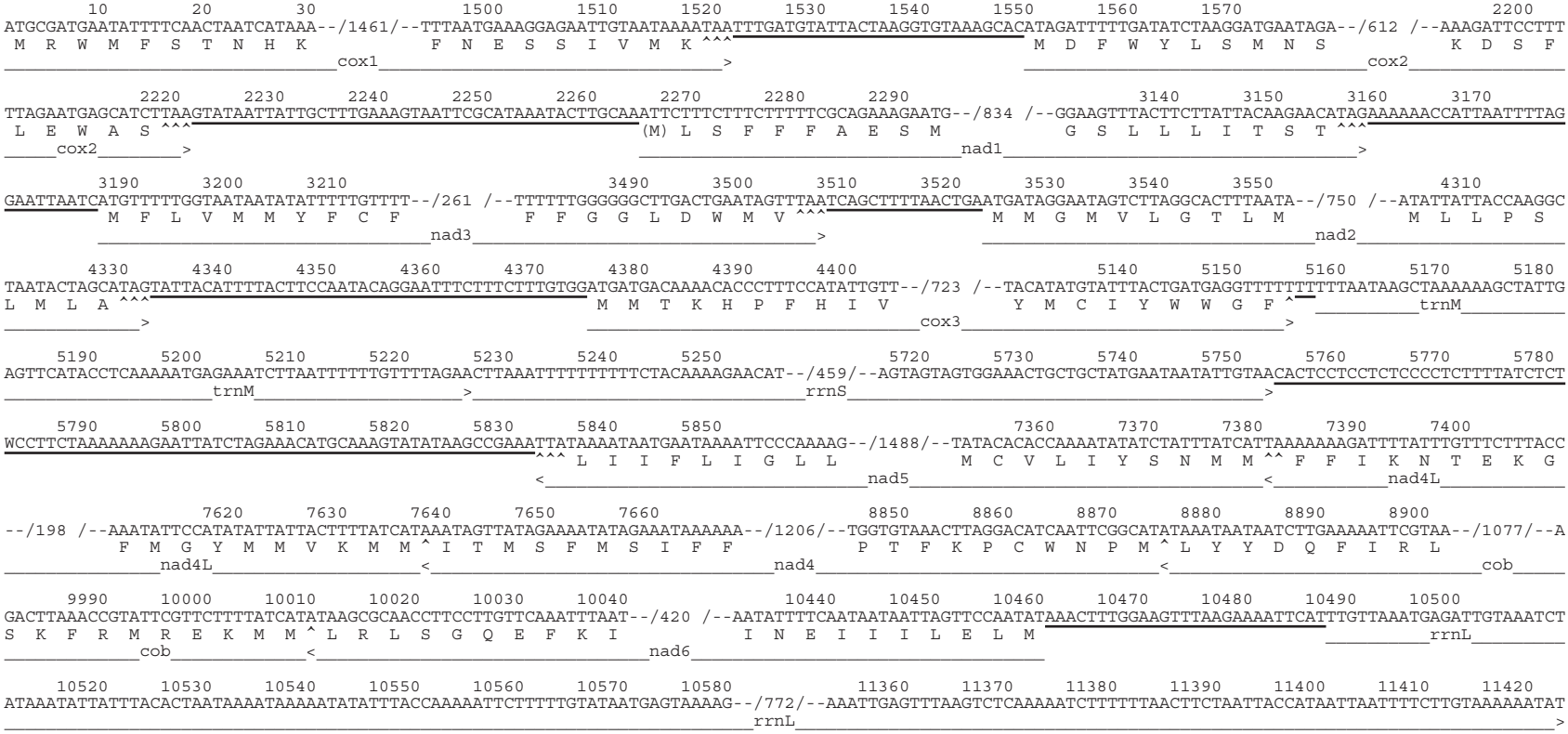
The single most parsimonious tree recovered from analysis of 2,785 aligned positions of inferred amino acids from eight mitochondrial protein genes (*cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad3*, *nad4*, and *nad5*). 1,802 positions are parsimony informative. Consistency index (CI) is 0.595 and retention index (RI) is 0.400. Numbers below branches are percent bootstrap support followed by Bremer support values. The latter indicates the number of evolutionary changes in the shortest tree that does not contain the node in question. Nodes with a bootstrap value of less than 90% (of which the highest was 50%) have been collapsed; these have a Bremer support in the range of 3-8.



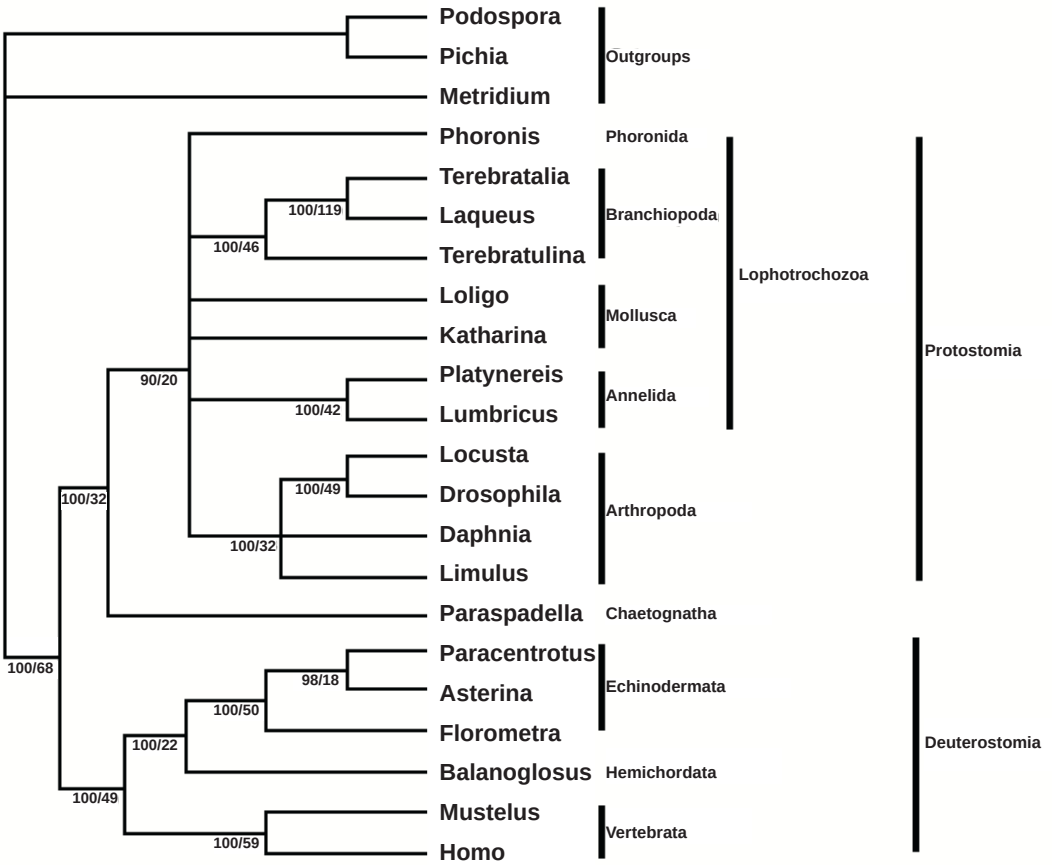
Helfenbein et al., Figure 1



Helfenbein et al. Figure 2



Helfenbein et al., Figure 3



Helfenbein et al., Figure 4