Comparison of Mitochondrial Genome Sequences of Cetartiodactyls (Mammalia, Laurasiatheria)

Wenceslao Parrilla-Martinez, Luis Manuel Guevara-Chumacero

Department of Biology, Autonomous Metropolitan University, Iztapalapa Campus, Mexico City, Mexico

Email address: lmgc@xanum.uam.mx (L. M. Guevara-Chumacero)

To cite this article: Wenceslao Parrilla-Martinez, Luis Manuel Guevara-Chumacero. Comparison of Mitochondrial Genome Sequences of Cetartiodactyls (Mammalia, Laurasiatheria). International Journal of Genetics and Genomics. Vol. 3, No. 3, 2015, pp. 26-31. doi: 10.11648/j.ijgg.20150303.11

Abstract: Comparing complete animal mitochondrial genome sequences is becoming increasingly common as a model for genome evolution and phylogenetic reconstruction. In the present work, we compare the complete mitochondrial genome sequences of five species of cetaceans and artiodactyls and infer phylogenetic relationships among them. The genome of the taxa contains the 37 genes found in a typical mammalian genome, a general structure that is highly-conserved among species. Phylogenetic trees constructed using MP, ML, and BI methods show a similar topology, and indicate the paraphyly of Artiodactyla, due to the sister-group relationship between the Cetacea and the Hippopotamidae. The study confirms that mitogenomics is a useful tool for research on mammal phylogenetics, but recognizes that increased taxon sampling is still required to resolve existing differences between nuclear and mitochondrial gene trees.

Keywords: Cetartiodactyls, Mitogenomics, Base Composition, Phylogenetic Relationship

1. Introduction

The sequence and structure of mitochondrial (mt) genomes are widely used to provide information on comparative and evolutionary genomics, molecular evolution, gene flow patterns, phylogenetics, and population genetics [1, 2]. Several recent analyses have demonstrated that complete mt genomes provide higher levels of support than those based on individual or partial mt genes [3, 4, 5]. The mammalian mitochondrial genome is a closed-circular, double-stranded DNA molecule about 16.5 kb in length that usually encodes genes for only thirteen protein products, two ribosomal RNAs, and 22 transfer RNAs [6].

Cetartiodactyla is one of the most diversified orders of mammals, with 332 extant species grouped into 132 genera. The members of this taxonomic group were originally divided into two different orders: Artiodactyla and Cetacea [7], but recent literature suggests a close relationship between these two orders based on a host of paleontological [8], morphological [9], and molecular [10, 11, 12] studies. Molecular data analyses, in particular, indicate that the Cetacea are sisters of the Hippopotamidae [10, 13], thus suggesting that all species of Artiodactyla and Cetacea belong to one sole order, called Cetartiodactyla [14]. This order includes all species of Cetacea, Hippopotamidae, Antilocapridae, Bovidae, Cervidae, Giraffidae, Moschidae, Tragulidae, Suidae, Tayassuidae, and Camelidae.

This study compared complete mitochondrial genomes from species of Artiodactyla and Cetacea in order to gain a better understanding of the molecular evolution of the mitochondrial genome and infer phylogenetic relationships among these groups.

2. Materials and Methods

Complete mitochondrial genomes were downloaded from GenBank and the heavy-strand encoded protein-coding genes were aligned according to [15]. The analyses based on complete mitochondrial genomes included five species: two cetaceans (Balaenoptera musculus - GenBank number X72204; Caperea marginata - GenBank number AP006475), and three artiodactyls (Hippopotamus amphibius - GenBank number AP003425; Moschus chrysogaster - GenBank number AP003427; Tayassu tajacu - GenBank number AP003427).
2.1. Comparative Analyses

Each one of the five mitochondrial protein-coding genes was compared. The initial and final codons were predicted for 13 protein-coding genes by comparing the amino acid sequences. Ribosomal RNA (rRNA) and transfer RNA (tRNA) genes were also identified by comparison with other complete mitochondrial rRNAs from cetartiodactyls. All comparisons were performed using CLC Genomics Workbench software version 8.0.1 (CLC bio).

2.2. Phylogenetic Analyses

The complete mitochondrial genomes were aligned using BioEdit [16], with subsequent visual confirmation. All those regions that showed ambiguity due to the position of the gaps were excluded from the analyses to avoid erroneous hypotheses of primary homology. Thus, the control region was not included in the alignment because of the presence of many ambiguous gaps. The final alignment consisted of 13,813 nucleotide sites (nt).

Phylogenetic reconstructions were obtained under Maximum Parsimony (MP) and Maximum Likelihood (ML) criteria with PAUP* 4.0b10 [17]. Under MP, trees were obtained considering gaps as a fifth state and after a heuristic search with an initial tree obtained via stepwise addition (random input order) of the taxa, followed by complete tree-bisection-reconnection (TBR) branch-swapping. The robustness of the topology was then assessed through bootstrapping [18] after 10,000 replicates. The most appropriate model of nucleotide evolution was selected using ModelTest 3.7 [19]. A heuristic ML search was then conducted using this model. Again, the robustness of the topology was evaluated using bootstrap resampling [18] after 10,000 replicates.

Bayesian phylogenetic analysis (BI) was also performed using MrBayes 3.1.2 [20]. The best-fit model (GTR+I+G) of sequence evolution for Bayesian analysis was obtained by ModelTest 3.7 [19] under the Akaike Information Criterion, with the following parameters: Nst = 6, rates = gamma, Ngen = 6,000,000, frequency = 100, chains = 5. After eliminating the first 20,000 trees as “burn-in,” we constructed a majority-rule consensus tree with Bayesian posterior probabilities for each node.

In accordance with [12], trees were rooted using sequences from Equus asinus (GenBank number NC_001788) and Equus caballus (GenBank number NC_001640).

3. Results and Discussion

3.1. Characteristics of the Mitochondrial Genome

The five mitochondrial genomes represent circular, double-stranded DNA molecules that ranged from 16,353 to 16,836 nucleotides and contained the 37 genes, including 22 tRNA genes, two rRNA genes (rrnS and rrnL), 13 protein-coding genes (COX1-3, NAD1-6, NAD4L, cytochrome b and ATP6,8), and a non-coding region (Table 1 and Figure 1). Of these genes, 8 tRNAs and ND6 are encoded by the L-strand, while all others are encoded by the H-strand. Arrangement is the same in all species of cetaceans and artiodactyls, and has been seen to be similar in other mammalian genomes, such as Artibeus jamaicensis bats [21] and Rhinolophus monoceros [22] that, according to [23], are closely related phylogenetically to cetaceans and artiodactyls.

All protein-coding genes of the mtDNA of the five species have a methionine start codon (ATR), except ND3 in H. amphibi us (ATT) and ND5 in M. chrysogaster. Also, all present TTA in ND6 and, with the exception of M. chrysogaster, all have GTC in ND4L. Most protein-coding genes appear to be terminated by TAR, though this stop codon is incomplete in various species, such as ND2, COI, ATP8, ND4L, and ND5.

Overlapping nucleotides between the mt genomes of the species examined in this study ranged from 1 to 45 base pairs (bp), (Table 1), the same as in other cetaceans, such as Orcinus orca [24] and artiodactyls like Bison bison [25].

The overall base composition of the mt genome shows no significant differences between the cetacean (B. musculus and C. marginata) and artiodactyl species (H. amphibi us, M. chrysogaster and T. tajacu, Figure 2). In fact, the greatest differences were found between species of artiodactyls. The percentages of nucleotides found were: C (27.6, 26.6 vs. 27.8, 28.6), G (13.0, 13.0 vs. 14.0, 12.9, 13.5%), A (24.6, 28.1, 25.3% vs. 24.6, 27.5 vs. 24.6, 28.1, 25.3%), and T (26.6, 27.5 vs. 28.6, 25.0, 27.0%).

Sequence analyses revealed the presence of two ribosomal RNA genes (rRNA): 12S rRNA and 16S rRNA. The size and arrangement of these genes concur with those of other cetaceans and artiodactyls. The sizes of the 12S rRNA and 16S rRNA were similar in all species, with a range of 951-974 and 1565-1576, respectively. The 12S rRNA gene is located between tRNAPhe and tRNA Val, while the 16S rRNA gene is between tRNA Val and tRNACys, and varied from 32 bp to 37 bp, similar to such bat species as Corynorhinus rafinesquii, Lasiusora borealis and Artibeus lituratus [27]. The sequence lengths of the control region, meanwhile, presented a slight variation in the five species, of 916-959 bp. This region exhibited higher AT than GC content, similar to other metazoans [28]. The highest amounts of GC were determined to be 41.2% for H. amphibi us and 38.6% for B. musculus.
| Name of gene | Location (Size in bp) | Artiodactyla | Strand | Start codon | Stop codon |
|-------------|-----------------------|--------------|---------|-------------|------------|
| tRNA-Phe    | 422-494 (73)          |              | H       | 1-70 (70)   |            |
| tRNA-Ile    | 4143-4211 (69)        |              | H       | 3693-3762 (70) |          |
| tRNA-Val    | 1468-1534 (67)        |              | H       | 3693-3762 (70) |          |
| tRNA-Trp    | 5394-5462 (69)        |              | H       | 4945-5012 (68) |          |
| tRNA-Gln    | 4209-4281 (73)        |              | L       | 3760-3832 (73) |          |
| tRNA-Met    | 4283-4351 (69)        |              | L       | 3833-3902 (70) |          |
| tRNA-Tyr    | 5712-5777 (66)        |              | L       | 5266-5333 (68) |          |
| tRNA-Asn    | 5538-5611 (74)        |              | L       | 5094-5167 (74) |          |
| tRNA-Cys    | 5644-5711 (68)        |              | L       | 5166 – 5199 (33) |       |
| tRNA-Lys    | 8156-8223 (68)        |              | H       | 7715-7781 (67) |          |
| ATP synthase F0 subunit 8 (ATP8) | 8225-8416 (192) |              | H       | 7783-7986 (204) |          |
| ATP synthase F0 subunit 6 (ATP6) | 8386-9066 (681) |              | H       | 7944-8624 (681) |          |
| Cytochrome c oxidase subunit III (COIII) | 9066-9851 (786) | 8647-9432 (786) | H | 8624-9407 (784) |          |
| tRNA-Gly    | 9851-9991 (69)        |              | H       | 9408-9475 (68) |          |
| NADH dehydrogenase subunit 3 (ND3) | 9920-10266 (347) | 9501-9847 (347) | H | 9477-9823 (347) |          |
| tRNA-Arg    | 10267-10335 (69)      |              | H       | 9824-9891 (68) |          |
| NADH dehydrogenase subunit 4L (ND4L) | 10336-10632 (297) | 9917-10213 (297) | H | 9892-10188 (297) |          |
| NADH dehydrogenase subunit 4 (ND4) | 10626—12003 (1378) | 10207-11584 (1378) | H | 10182-11559 (1378) |          |

Table 1. Characteristics of the mitochondrial genome in two species of cetaceans and three species of artiodactyls.
| Name of gene | Location (Size in bp) | Cetacea | Artiodactyla | Strand | Start codon | Stop codon |
|--------------|-----------------------|---------|--------------|---------|-------------|------------|
|              |                       | Balaenoptera musculus | Caperea marginata | Hippopotamus amphibius | Moschus chrysogaster | Tayassu tajacu |
| tRNA-His     | 12004-12072 (69)      | 11585-11653 (69)   | 11564-11633 (70) | 11547-11616 (70) | 11560-11627 (68) | H           |
| tRNA-Ser (AGY) | 12073-12133 (61)   | 11654-11714 (61)   | 11634-11692 (59) | 11617-11676 (60) | 11686 (59) | H           |
| tRNA-Leu (CUN) | 12135-12204 (70)  | 11716-11785 (70)   | 11694-11763 (70) | 11678-11747 (70) | 11687-11756 (70) | H           |
| NADH dehydrogenase subunit 5 (ND5) | 12205-14025 (1821) | 11786-13606 (1821) | 11764-13584 (1821) | 11748-13568 (1821) | 11757-13577 (1821) | H, ATA<sup>a,b,c</sup> ATTA<sup>d</sup> |
| NADH dehydrogenase subunit 6 (ND6) | 14009-14536 (528) | 13590-14117 (528) | 13568-14098 (531) | 13552-14079 (528) | 13556-14058 (528) | L, ATG TAA |
| tRNA-Glu     | 14537-14605 (69)     | 14118-14186 (69)   | 14096-14164 (69) | 14080-14148 (69) | 14089-14157 (69) | L           |
| Cytochrome b (Cytb) | 14610-15749 (1140) | 14191-15330 (1140) | 14169-15308 (1140) | 14153-15292 (1140) | 14162-15301 (1140) | H, ATG AGA |
| tRNA-Thr     | 15750-15821 (72)     | 15331-15402 (72)   | 15309-15379 (71) | 15296-15365 (70) | 15302-15371 (70) | H           |
| tRNA-Pro     | 15821-15887 (67)     | 15402-15468 (67)   | 15378-15443 (66) | 15365-15430 (66) | 15371-15435 (65) | L           |
| D-loop       | 15888-16402, 1-421 (935) | 15469-16384 (916) | 15444-16402 (959) | 15431-16353 (923) | -           |

**Figure 1.** Comparative mitogenomics between cetacean and artiodactyl species. In blue arrows shown the 13 essential respiratory chain sub-units: ND1–ND6 (NADH dehydrogenase sub-unit 1-6) and ND4L (NADH dehydrogenase sub-unit 4L gene), CYT B (cytochrome b gene), COX 1-COX 3 (cytochrome c oxidase 1-3), and the ATPase6 and ATPase8 genes. Also shown are the two ribosomal RNA genes (12S rRNA and 16S rRNA) (red arrow), and the 22 transfer RNA genes (red arrow). Transfer RNA genes are Phe, Val, Leu, Ile, Glh, Met, Trp, Ala, Asn, Cys, Tyr, Ser, Asp, Lys, Gly, Arg, His, Glu, Thr, and Pro.
3.2. Phylogeny

Maximum Parsimony, Maximum Likelihood and Bayesian phylogenetic analyses based on an alignment of 13,813 bp resulted in a robust tree (Figure 2) with high bootstrap values that support (100%) interspecific relationships. Results indicate the paraphyly of Artiodactyla due to the sister-group relationship between Cetacea and Hippopotamidae. According to [12], these relationships formed the Whippomorpha group. Moreover, they are in full agreement with recent works that suggest a close assembly between Artiodactyla and Cetacea based on a host of paleontological [8], morphological [9], and molecular [10, 11, 12] studies. To render taxonomy compatible with this phylogeny, Montgelard et al. [14] proposed placing all species of Artiodactyla and Cetacea in the same order, called Cetartiodactyla.

Genetic distances based on the evolutionary model selected by ModelTest were lower among the species of cetaceans (12.1%) than among the artiodactyls (22.5-23.3%). These latter values were similar to those seen in the between-taxon comparisons of the two orders (22.0-24.8%). These results provide a reasonable basis of support for mitogenomic relationships within Cetartiodactyla. Although anatomists had strongly claimed the monophyly of Artiodactyla for 150 years, it has been identified in all extant terrestrial cetartiodactyls and Eocene whales, a common ancestor of Cetartiodactyla that exhibits a double-pulley astragalus and a paraxonic foot [29].

4. Conclusions

This study presents the first comparison of entire mitochondrial genomes for various cetaceans and artiodactyls while also providing support for the notion that mitogenomics is a powerful tool for investigating genome evolution and phylogenetic relationships in mammals and suggesting that the general structure among species of cetaceans and artiodactyls is highly-conserved. Finally, the comparison of phylogenetic results requires the use of biparental inherited nuclear markers (i.e., introns).

Acknowledgment

The authors are grateful for the administrative and logistical support of the Department of Biology, Universidad Autónoma Metropolitana, Iztapalapa Campus (Mexico City). Daniel E. Amaro Sánchez and Paul C. Kersey Johnson provided insightful comments and suggestions that helped improve this manuscript.

References

[1] Wilson K., Cahill V., Ballment E., Benzie J. (2000) The complete sequence of the mitochondrial genome of the crustacean Penaeus monodon: are malacostracan crustaceans more closely related to insects than to branchiopods? Mol Biol Evol 17:863–874

[2] Salvato P., Simonato M., Battisti A., Negrisolo E. (2008) The complete mitochondrial genome of the bag-shelter moth Ochrogaster lunifer (Lepidoptera, Notodontidae). BMC Genomics 9:331

[3] Cummings MP., Otto SP., Wakeley J. (1995) Sampling properties of DNA sequence data in phylogenetic analysis. Mol Biol Evol 12: 814–822
[4] Russo CA, Takezaki N, Nei M (1996) Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol Biol Evol 13:525–536

[5] Zardoya R, Meyer A (1996) Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. Mol Biol Evol 13:933–942

[6] Taaman JW (1999) The mitochondrial genome: structure, transcription, translation and replication. BBA Bioenergetics 1410:103–123

[7] Wilson DE, Reeder DM (2005) Mammal species of the world. A taxonomic and geographic reference, 3rd. ed., Johns Hopkins University Press

[8] Thewissen JG, Hussain ST (1993) Origin of underwater hearing in whales. Nature 361:444–445

[9] Thewissen JG, Williams EM, Roe LJ, Hussain ST (2001) Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. Nature 413:277–281

[10] Murphy DA, Stein JA, Schlinger W, Maibach E (2001) Conceptualizing the multidimensional nature of self-efficacy: assessment of situational context and level of behavioral challenge to maintain safer sex. Health Psychology 20:281–290

[11] Hassanin A, Delsuc F, Roipiquet A, Hammere C, Vuuren BJ, Matthee C, Ruiz Garcia M, Catzeflis F, Areskoug V, Couloux A (2012) Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. Comptes Rendus Biologies 335:32–50

[12] Wang Q, Yang C (2013) The Phylogeny of the Cetartiodactyla Based on Complete Mitochondrial Genomes. Intern J Biol 5:30–36

[13] Nikaido M, Rooney AP, Okada N (1999) Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc Nat Acad Sci USA 96:10261–10266

[14] Montgierlad C, Catzeflis FM, Douzery E (1997) Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. Mol Biol Evol 14:550–559

[15] Nikaido M, Kawai K, Cao Y, Harada M, Tomita S, Okada N, Hasegawa M (2001) Maximum likelihood analysis of the complete mitochondrial genomes of eutherians and a reevaluation of the phylogeny of bats and insectivores. J Mol Evol 53:508–516

[16] Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–8

[17] Swofford DL (2002) PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods) 4.0 edn Sinauer Sunderland MA

[18] Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

[19] Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics Applications Note 14:817–818

[20] Ronquist F, Haelstenbeck JP (2003) Mrbayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

[21] Pumo DE, Finamore PS, Franke WR, Phillips CJ, Tarzami S, Balzarano D (1998) Complete mitochondrial genome of a neotropical fruit bat, Artibeus jamaicensis, and a new hypothesis of the relationships of bats to other eutherian mammals. J Mol Evol 47:709–717

[22] Lin YH, McLenachan PA, Gore AR, Phillips MJ, Ota R, Hendy MD, Penny D (2002) Four new mitochondrial genomes and the increased stability of evolutionary trees of mammals from improved taxon sampling. Mol Biol Evol 19:2060–2070

[23] Meredith RW, Janecka JE, Gatyse J, Ryder OA, Fisher CA, Teeling EC, Minos TLL, Staderl T, Rabosky DL, Honeycut RL, Flynn JJ, Inram CG, Steiner C, Williams TL, Robinson TJ, Burk-Herrick A, Westerman M, Ayoub NA, Springer MS, Murphy WJ (2011) Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science 334:521–524

[24] Morin PA, Archer FL, Foote AD, Vilstrup J, Allen EE, Wade P, Durban J, Parsons K, Pitman R, Li L, Bouffard P, Abel Nielsen SC, Rasmussen M, Willerslev E, Gilbert MTP, Harks T (2010) Complete mitochondrial genome phylogenetic analysis of killer whales (Orcinus Orca) indicates multiple species. Genome Research 20:908–916

[25] Douglas KC, Halbert ND, Kolenda C, Childers C, Hunter DL, Derr JN (2011) Complete mitochondrial DNA sequence analysis of Bison bison and bison–cattle hybrids: Function and phylogeny. Mammalion 11:166–175

[26] Coskun PE, Ruiz-Pesini E, Wallace DC (2003) Control region mtDNA variants: longevity, climatic adaptation, and a forensic conundrum. Proc Nat Acad Sci USA 100:2174–21726

[27] Meganathan PR, Pagan HJ, McCulloch ES, Stevens RD, Ray DA (2012) Complete mitochondrial genome sequences of three bats species and whole genome mitochondrial analyses reveal patterns of codon bias and lend support to a basal split in Chiroptera. Gene 492:121-129

[28] Asakawa S, Kumazawa Y, Araki T, Hinemo H, Miura KI, Watanabe K (1991) Strand-specific nucleotide composition bias in echinoderms and vertebrate mitochondrial genomes. J Mol Evol 32: 511–520

[29] Gingerich PD, ul Haq M, Zalmout IS, Khan IH, Malkani MS (2001) Origin of whales from early artiodactyls: hands and feet of Eocene Protocetidae from Pakistan. Science 293:2239–2242