

Molecular phylogenetics, barcoding and, speciation genetics: A perspective from mitogenomics

Yuri Ph. Kartavtsev

Abstract

Mitogenome is described for catfish, *Liobagrus obesus* and compared with other fish and vertebrates. For the protein-coding genes a nucleotide bias revealed. Such bias in favor of pyrimidine content is well known and now it is substantiated statistically in bulk of fish species. The properties of mitogenome of vertebrate animals will be briefly overviewed. Nucleotide diversity for the genes Cyt-b and Co-I was analyzed with special reference to the genetics of speciation. Genetic divergence of populations (1) and taxa of different rank, such as subspecies, semispecies and sibling species (2), species (3), species from different genera within a family (4), and species from separate families within an order (5) have been compared on p-distances. Data for 20,731 vertebrate and invertebrate animal species reveal various and increasing levels of divergence of the sequences of two genes in the five groups compared. Mean unweighted scores of p-distances (%) for five groups are: Cyt-b (1) 1.38 ± 0.30 , (2) 5.10 ± 0.91 , (3) 10.31 ± 0.93 , (4) 17.86 ± 1.36 , (5) 26.36 ± 3.88 and Co-I (1) 0.89 ± 0.16 , (2) 3.78 ± 1.18 , (3) 11.06 ± 0.53 , (4) 16.60 ± 0.69 , (5) 20.57 ± 0.40 . The results of the analysis of the nucleotide divergence within species and higher taxa of animals suggest that a phyletic evolution in animals is likely to prevail at the molecular level, and speciation mainly corresponds to the geographic mode. The approach suggested that allows recognize the speciation modes formally with the operational genetic criteria. Such approach may help to solve a key problem of the biological species concept, i.e. define species without knowledge on the reproductive isolation.

Fifteen years have passed since the publication of a novel approach for sequencing the whole mitochondrial genome (mitogenome) of fishes (ca. 16,500 bp) using a combination of the long PCR technique (Cheng et al. 1994) and a number of fish-versatile primers (Miya and Nishida 1999). During this period (1999–2014), our research group has assembled nucleotide sequences for ≈ 1340 whole mitogenomes from fishes using this method (including unpublished ones) and published 83 peer-reviewed papers on mitochondrial genomics (mitogenomics) of fishes (including those papers based on partial mitogenome sequences $>2,000$ bp; Fig. 1). Most of those papers (77 papers) are more or less phylogeny oriented with the exception of five reports of single mitogenomes from commercially or zoologically important species (e.g., Japanese sardine, Inoue et al. 2000b), a paper addressing the highly conservative gene arrangement in vertebrate mitogenomes based on a comparative genomic approach (Satoh et al. 2010),

and a paper describing a database specifically designed for fish mitogenomes (MitoFish) with a novel gene annotation pipeline. Apparently, the course of this direction towards phylogenetics is due to a successful demonstration of the utility of the mitogenomic data for resolving higher-level relationships of fishes at an earlier stage of the study. Consequently, the “mitogenomic era” was and an unprecedented number of contributions to molecular phylogenetics and evolution of fishes have been published from this single research group. Although members of this research group comprise mainly Japanese professional scientists, postdocs and graduate students, the affiliations of the 151 coauthors of these papers are distributed across 17 different countries, representing truly international collaborative efforts.

Taxonomic coverage of those studies is extensive encompassing three of the four major fish lineages from chondrichthyans (e.g., Inoue et al. 2010b; Aschliman et al. 2012) to various actinopterygians (e.g., Inoue et al. 2003a; Ishiguro et al. 2003; Miya et al. 2003; Saitoh et al. 2003) and even to sarcopterygians (coelacanth, Inoue et al. 2005). Notably, our research group has never performed *de novo* sequencing for mitogenomes from agnathans; however, Inoue et al. (2010b) used two agnathans as outgroups in their study of holocephalan evolutionary history. Thus, the temporal dimension of our studies is well over 450 million years, considering that a common ancestor of those four vertebrate lineages emerged in the Paleozoic (Benton et al. 2009). In terms of the spatial dimensions, our phylogenetic studies cover the global habitat diversity, including freshwaters on major continental landmasses coral reefs in tropical and subtropical watersdemersal waters (e.g., Macrouridae, see Satoh et al. 2006), and combinations across those habitats.

As of 30 September 2014, those 83 mitogenomic papers have been cited 5,303 times in total, with the most frequently cited paper reaching 543 citations, followed by 14 papers with ≥ 100 citations. A number of the mitogenomic papers have been cited in standard textbooks in ichthyology, such as “Fishes of the World” and “The Diversity of Fishes” (Helfman et al. 2009), and even in textbooks of other disciplines, such as “Vertebrate Palaeontology” and “Evolution” In the 4th edition of “Fishes of the World,” mentioned our research activities, stating “... Some of the new, exciting molecular work is being done in the laboratories of ... Masaki Miya in Chiba, Japan; and Mutsumi Nishidain Tokyo, Japan. These workers and

their colleagues as well as many others are finding support for new clades that must be taken seriously, especially in those areas where agreement is found when different approaches are taken.” In addition to those citations in the textbooks, our article demonstrating a sister group relationship between pike and salmon was featured in a review article of Trends in Ecology & Evolution for discussing the evolutionary origin of anadromous migrations of salmon. Our recent discovery of “three deep-fish families into one” was highlighted in a “News & Views” section in Nature and our discovery of deep-ocean origin of the freshwater eels was highlighted in a “Feature” section in Current Biology.

Other than those scientific impacts, some of our recent papers have attracted considerable attention from various media outlets, such as popular science magazines, newspapers, radio, and television.

This work is partly presented at 2nd International Conference on Big Data Analysis and Data Mining 30-December 01, 2015 San Antonio, USA

Yuri Ph. Kartavtsev,
A.V. Zhirmunsky Institute of Marine Biology, Russia, E-mail: yuri.kartavtsev48@hotmail.com