Unique mitochondrial gene order in *Xenodermichthys copei* (Alepocephalidae: Otocephala) – a first observation of a large-scale rearranged 16S–WANCY region in vertebrates

Jan Yde Poulsen,a,b Jan Yde Poulsen, Tetsuya Sado and Masaki Miya

aDepartment of Fish and Shellfish, Pinggortitaleriffik (Greenland Institute of Natural Resources), Nuuk, Greenland; bDepartment of Ecology and Environmental Sciences, Natural History Museum and Institute, Chiba, Japan

**ABSTRACT**

Mitochondrial gene orders in vertebrates have proven to comprise strong phylogenetic characters when reconstructing phylogenetic relationships. Unfortunately, larger gene-order rearrangements, compared to the canonical mitochondrial gene order found in most vertebrates, have been difficult to determine using conventional (polymerase chain reaction) PCR amplification and automated sequencing. We here present the complete mitogenomic DNA sequence of the blunt snout smooth-head *Xenodermichthys copei* of the family Alepocephalidae – a species that previously failed determination of its mitogenomic DNA sequence from non-next generation sequencing methods. Long PCR amplification and subsequent Illumina MiSeq sequencing of the long amplified fragments proved successful in determining the mitogenome of *X. copei*. The mitochondrial gene order, including many INC-regions, is unique among vertebrates due to a large rearranged segment in between the 12S rRNA and the Cox1 DNA gene sequences. The segment includes a rearranged WANCY region (WNCAY), a large control region, relocated Leu1-, Ile- and Met-tRNA gene sequences, relocated 16S rRNA and ND1 DNA gene sequence including multiple INC-regions. *Xenodermichthys* is one of only four genera of fishes in the family Alepocephalidae that show bioluminescent photophores present – a group of slickheads that have proven difficult in determining their complete mitogenomic DNA sequences. Gene orders are expected to resolve long-standing taxonomic and phylogenetic questions pertaining to these genera.

**INTRODUCTION**

Rearranged mitochondrial (mt) gene orders in vertebrates have proven to be solid phylogenetic characters, due to the fact that most vertebrates show a highly conservative gene order. The typical mt genome of vertebrates consists of approximately 16,500 base pairs with only few non-coding base pairs in between coding DNA gene sequences (Boore and Brown 1998). Gene order rearrangements have been shown to happen from either slipped-strand mispairing during replication (Levinson and Gutman 1987) or by illicit priming of replication by tRNAs (Cantatore et al. 1987). Regardless of any mechanisms involved in the replication processes, the mt genome must be considered highly economical, and changes to this scheme become informative on even small scales, such as e.g. rearrangement of shorter fragments involving coding tRNA gene sequences or longer protein-coding DNA sequences. Interestingly, rearrangement events have occurred a number of times across multiple independent vertebrate lineages, including within teleost fishes (e.g. Inoue et al. 2003; Poulsen et al. 2013; 2010b), reptiles (e.g. Kumazawa et al. 1996; 2014), amphibians (e.g. Kurabayashi et al. 2008; Xia et al. 2014) and birds/crocodiles (e.g. Mindell et al. 1998; Li et al. 2007). Mitochondrial gene order rearrangements have not been observed in mammals (e.g. Arnason et al. 2008) or chondrichthyans (e.g. Inoue et al. 2010) although they are present in agnathans (Delarbre et al. 2000), all hemichordates (e.g. Singh et al. 2009), and cephalochordates (e.g. Nohara et al. 2005). There is little doubt concerning the phylogenetic utility of mt gene orders, although ‘hotspots’ for duplications are present and must be considered in the context of the gene orders observed (San Mauro et al. 2006).

We here report on the complete mitogenomic DNA sequence and mt gene order of the blunt snout smooth-head *Xenodermichthys copei* (Gill 1884) that previously failed in determination when using long amplification (LA) polymerase chain reaction (PCR), nested PCR and Sanger sequencing (Lavoué et al. 2008). We have employed LA PCR and the Illumina MiSeq NGS (Next Generation Sequencing) platform, thereby initiating mitogenomic determinations of mt gene orders in photophore-possessing slickheads (Alepocephalidae). Extensive laboratory works in the past have shown the determination of mt genomes in photophore-possessing alepocephalids a problematic issue compared to other alepocephaliforms and fish groups in general.
Material and methods

The X. copei specimen determined for its complete mitogenomic DNA sequence was caught off southeast Greenland (61.56°N, 39.52°W) on 31 July 2013, at 1367 m fishing depth, onboard R/V Pamiut (Greenland Institute of Natural Resources). It has the tissue/field number JYP#9819, and is vouched at the Natural History Collections in Bergen, Norway, as ZMUB 21822 (collected and identified by J.Y. Poulsen). The specimen was initially Cox1 barcoded as part of the Greenland Fishes Barcoding Project (Poulsen et al. 2018) and deposited as GLF087 at the Barcoding of Life Database (BOLD, Ratnasingham and Hebert 2007).

Subsequently, LA PCR was performed according to Cheng et al. (1994) with universal fish primers, which have been successful also with slickheads (Miya and Nishida 1999; Byrkjedal et al. 2011). The LA fragments were pooled and the concentration adjusted for Illumina Nextera XT Library preparation that uses transposons for tagging and fragmentation. The DNA library was sequenced using the Illumina MiSeq Benchtop Sequencher platform (San Diego, CA) at the Chiba Natural History Museum and Institute (Chiba, Japan). The mitogenomic DNA sequence was assembled using MITObim (Hahn et al. 2013), using the previously determined Cox1 barcode as seed, and annotated using Sequencher version 5.0.1 (Genecodes, Ann Arbor, MI) and tRNA-Scan-SE version 2.0 (Lowe and Eddy 1997) software, paying special attention to H- and L-stands in relation to presence of tRNA gene sequences. The Control Region (CR) of X. copei was scrutinized for conserved motifs according to various fishes (Lee et al. 1995; Brzuzan and Ciesielski 2002; Rocha-Olivares et al. 2005). However, as no closely related taxa have been identified for the CR in the Alepocephalidae for comparisons, and CR comparisons to other slickheads proved difficult, deliminations of conserved motifs must be considered putative. The newly determined mt genome was deposited as AP017982 in the DNA Databank of Japan, European Molecular Biology Laboratory, and GenBank repositories.

Results

The complete mitogenomic DNA sequence of X. copei (ZMUB 21822) is 18,235 base pairs including 13 protein-coding genes, 2 tRNA genes, and 22 tRNA genes as found in other vertebrates. The gene order is, on the contrary, unique among vertebrates showing a 12S-V-INC-L1-ND2-Q-ND1-I-Cox1 gene order (INC = Intergenic Non-Coding region). The canonical mt gene order in vertebrates is shown in Figure 1(A), mt gene order of X. copei is shown in Figure 1(B) and the elongated CR consisting of 1310 base pairs positioned between tRNAs-Gln and tRNAs-Ile is shown in Figure 1(C). Abbreviations follow standard mitochondrial nomenclature not to be repeated here. The CR of X. copei shows some motifs present which are usually observed in fishes, although here noted with some reservations; four Termination Associated Sequences (-TACAT-) denoted as TAS 1–4; six complementary TAS (-ATGTA-) denoted as cTAS 1–6; two Conserved Sequence Blocks denoted as CSB D, and CSB 2, as CSB 1 and 3 could not be located and were therefore omitted; a binding site for the mt transcription factor (TF) with the sequence -CACCAGC-; a T-homopolymer of 11 tyrosine bases and a tentative poly-pyrimidine tract (-TTTTAACCCCCCCCCT-) denoted as PP (tentative because the segment is interspersed by two adenine bases).

Discussion

Mitogenome of Xenodermichthys copei

The mt genome of X. copei is unique as it is showing a gene order never before observed in vertebrates (Figure 1). As no other gene order rearrangements have been found in alepocephalids (Lavoué et al. 2008; Poulsen et al. 2009), the rearrangement event(s) causing the observed gene order is unknown. Transposable elements during replication could theoretically be involved, due to the reversed positions of the ND1 and 16S genes in X. copei (compared to the canonical vertebrate mt genome) including several juxta-positioned tRNA genes (Figure 1). However, considering the WNCAY tRNA gene order (as opposed to the canonical WANCY usually observed in vertebrates), this is not the parsimonious explanation at present. Instead, a large duplication event involving the region from 12S to WANCY could have taken place at one point, with subsequent deletions of the original gene sequences. This would have resulted in the newly observed gene order. The large INC-region observed between tRNA-Gln and 16S, including the tRNA-Ile gene, shows some sequence similarity to the control region of other slickheads. Considering the relatively short CR in the canonical position between tRNAs Pro and Phe (Figure 1(B)), the repositioned CR between tRNA-Gln and 16S is likely functional as several conservative motif sequences are found such as the CSB and TAS (Brown et al. 1986; Saccone et al. 1991). Tandem repeats are found throughout the CR and the number of base pair repeats are likely to be caused by nucleotide mispairing during replication (Lee et al. 1995). Slipped strand mispairing is also the most commonly cited hypothesis of gene duplications and subsequent deletions (Moritz and Brown 1986). The competing theory explaining mt gene order rearrangements involves ilicit priming of replication by tRNAs, resulting in tRNA genes associated with the CR (Cantatore et al. 1987; Jacobs et al. 1989). All current observations on gene-order rearrangements in fishes are most parsimoniously explained by duplications of parts of the mt genome with subsequent and possibly random deletions (e.g. Inoue et al. 2003). Examination of a previously determined non-complete mt genome of X. copei (Lavoué et al. 2008) showed a putative similar gene order as reported here, although the rearranged region in the latter study could not be determined using LA PCR and automated sequencing.

Gene order rearrangements are present in the closely related Platytroctidae (Lavoué et al. 2008; Poulsen et al. 2009) although all current phylogenetic hypotheses render the gene order rearrangements within the Alepocephalidae and Platytroctidae to be independent molecular events (Kawato et al. unpublished data). Gene order rearrangements in slickheads appear to be present only in photophore-possessing
taxa – a personal observation from years of laboratory work on almost all genera in the Alepocephalidae family. While we successfully determined the complete mitogenic DNA sequence of *X. copei* with LA PCR followed by MiSeq Sequencing, the LA PCR failed in the closely related genus Rouleina. Poulsen (2018a) showed *Xenodermichthys* and *Rouleina* closely related from barcoding Cox1 gene fragments, a result corroborated by morphology (Markle 1976), and we suspect that similar or larger scale rearrangements are present in *Rouleina* taxa thereby failing the LA PCR. Variable Number Tandem Repeats (VNTRs) usually observed in the 3 end of the CR in other fishes (e.g. Brzuzan and Ciesielski 2002) seems to be present throughout the entire CR in *X. copei*. We found no apparent structure in the positions of the VNTR observed. A single T-homopolymer is observed in the CR of *X. copei*, a DNA-repeat sequence known to have caused difficulties during the automated sequencing process (Poulsen et al. 2013; Satoh et al. 2016). As noted, laboratory work resulting in the studies Lavoué et al. (2008) and Poulsen et al. (2009) failed to determine complete mitogenic DNA sequences of both *Xenodermichthys* and *Rouleina* taxa. Concerning *X. copei*, the latter studies failed to determine the region ND1 to 16S during nested PCR on LA fragments, corresponding to a segment of the rearrangement observed in the complete mt genome of *X. copei* (ZMUB 21822) determined in this study (Figure 1(A)). Using NGS technologies, such as the MiSeq sequencing platform, facilitate the determination of complicated mitogenomes, although LA PCR is still an obstacle for some taxa regardless of tissue quality and method of sequencing.

Acknowledgements

*R/V Pomst* crew and staff (Greenland Institute of Natural Resources, Nuuk, Greenland) for help with sampling and taxonomic work onboard, I. Byrkjedal (Natural History Collections Bergen, Norway) for curating and providing information whenever needed, and I. Chemshirova (Zoological Society London, UK) for proofreading and correcting manuscript.

Disclosure statement

The authors report no conflicts of interest.

ORCID

Jan Yde Poulsen  
<http://orcid.org/0000-0003-2137-0759>
Arnason U, Adegoke JA, Gullberg A, Harley EH, Janke A, Kullberg M. 2008. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. Genres. 42:137–51. Doi:10.1016/j.gene.2008.05.024
Boore JL, Brown WM. 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Curr Opin Genet Dev. 8:668–674.
Brown GG, Gadaleta G, Pepe G, Saccone C, Sbisà E. 1986. Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. J Mol Biol. 192:503–511.
Bruzan P, Ciesielski S. 2002. Sequence and structural characteristics of mtDNA control regions of three coregonine species (Coregonus albula, C. lavaretus, and C. peled). Arch Hydrobiol Spec Issues Adv Linnol. 57: 11–20.
Byrkjedal I, Poulsen JY, Galbraith J. 2011. Leptodorma macrphalhalm n. sp., a new species of smooth-head (Otocephala: Alepocephalidae) from the Mid Atlantic Ridge. Zootaxa. 2876:49–56.
Cantatore P, Gadaleta MN, Roberti M, Saccone C, Wilson AC. 2018. Duplication and remodeling of tRNA genes during the evolution of mitochondrial genomes. Nature. 329:835–855. Doi:10.1038/s229536a0
Cheng S, Higuchi R, Stoneking M. 1994. Complete mitochondrial genome amplification. Nat Genet. 7:350–351.
Delarbre C, Escriva H, Gallut C, Barriel V, Kourilsky P, Janvier P, Laudet V, Cheng S, Higuchi R, Stoneking M. 1994. Complete mitochondrial genome of the agnathan Lampeta fluviatilis: bearings on the phylogeny of cyclostomes. Mol Biol Evol. 17:519–529.
Gill TN. 1884. Three new families of fishes added to the deep-sea fauna of the subarctic North Atlantic Ocean including three new distributional records and multiple molecular OTUs of Notacanthus cf. chemnitzii. Mar Biodivers. 48:1009–1025.
Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nuclic Acid Res. 41(13):e129.
Inoue JG, Miya M, Nishida M, Knudsen SW, Nishida M. 2009. Monophyly, and support for the new chordate phylogeny. BMC Genom. 10:534.
Jacobs HT, Asakawa S, Araki T, Miura K, Smith MJ, Watanabe K. 1989. Conservation of tRNA gene cluster in starfish mitochondrial DNA. Curr Genet. 15:193–206.
Kawato M, Poulsen JY, Tsuchida S, Ida H, Chikaraishi Y, Okouchi N, Oguri K, Goto S, Ozawa G, Tanaka S, et al. (unpublished data). Discovery of a colossal slickhead (Alepocephaliformes: Alepocephalidae): an active top-predator in the deep Suruga Bay. Sci Rep.
Kumazawa Y, Miya K, Ozawa T. 1996. Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. Mol Biol Evol. 13:1242–1254.
Kurabayashi A, Sumida M, Yonekawa H, Glaw F, Vences M, Hasegawa M. 2008. Phylogeny, recombination, and mechanisms of stepwise mitochondrial genome reorganization in mantellid frogs from Madagascar. Mol Biol Evol. 25:874–891.
Lavoué S, Miya M, Poulsen JY, Moller PR, Nishida M. 2008. Monophyly, phylogenetic position and inter-familial relationships of the Alepocephaliformes (Teleostei) based on whole mitogenome sequences. Mol Phylogenet Evol. 47:1111–1121.
Lee WJ, Conroy J, Howell WH, Kocher TD. 1995. Structure and evolution of teleost mitochondrial control regions. J Mol Evol. 41:54–66.
Levinson G, Gutman GA. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol. 4:203–221.
Li Y, Wu X, Ji X, Yan P, Amato G. 2007. The complete mitochondrial genome of salt-water crocodile (Crocodylus porosus) and phylogeny of crocodylians. J Genet Genom. 34:119–128.
Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:565–964.
Markle DF. 1976. Preliminary studies on the systematics of deep-sea Alepocephaloidea (Pisces: Salmoniformes). [PhD Thesis]. Virginia: The College of William and Mary.
Mindell DP, Sorensen MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. Proc Natl Acad Sci USA. 95: 10693–10697.
Miya M, Nishida M. 1999. Organization of the mitochondrial genome of a deep-sea fish, Gonostoma gracile (Teleostei: Stomiforimes): first example of transfer RNA gene rearrangements in bony fishes. Mar Biotecnol. 1:416–426.
Moritz C, Brown WM. 1986. Tandem duplications of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. Science. 233: 1425–1427.
Nohara M, Miya M, Nishikawa T. 2005. Evolution of the mitochondrial genome in Cephalochordata as inferred from complete nucleotide sequences from two Epigonichthys species. J Mol Evol. 60: 526–537.
Poulsen JY, Byrkjedal I, Willansen E, Rees D, Takeshima H, Satoth TP, Shinhohara G, Nishida M, Miya M. 2013. Mitogenomic sequences and evidence from unique gene rearrangements corroborate evolutionary relationships of Myctophiformes (Neoteleostei). BMC Biol. 13:111.
Poulsen JY, Miller MJ, Lavoué S, Knudsen SW, Nishida M, Miya M. 2009. Higher and lower-level relationships of the deep-sea fish order Alepocephaliformes (Teleostei: Otocephala) inferred from whole mitogenome sequences. Biol J Linn Soc. 98:923–936.
Poulsen JY. 2018a. New observations and ontogenetic transformation of photogenic tissues in the tubeshoulder Sagamichthys schnakenbecki (Platytretioidae, Alepocephaliformes). J Fish Biol.
Poulsen JY, Miller MJ, Sado T, Hanel R, Tsukamoto K, Miya M. 2018b. Resolving deep-sea pelagic sarcophangyromiform eel mysteries: identification of neocyema and monognathidae leptocephali and establishment of a new fish family “Neocyematidae” based on larvae, adults and mitogenomic gene orders. PLoS One. 13:e0199982.
Poulsen JY, Thorkslien S, Arboe NH. 2017. Identification keys to halosaurus and notacanthids (Notacanthiformes, Elopomorpha) in the subarctic North Atlantic Ocean including three new distributional records and multiple molecular OTUs of Notacanthus cf. chemnitzii. Mar Biodivers. 48:1009–1025.
Ratnasingham S, Hebert PDN. 2007. BOLD: the barcode of life data system (http://www.barcodinglife.org). Mol Ecol Notes. 7:355–364.
Rocha-Olivares A, Garber NM, Garber AF, Stuck KC. 2005. Structure of the mitochondrial genome in Cephalochordata as inferred from complete genome sequences. Biol J Linn Soc. 98:923.
San Mauro D, Gower DJ, Zardoya R, Wilkinson M. 2006. A hotspot of higher and lower-level relationships of the deep-sea fish order Alepocephaliformes (Teleostei: Otocephala) inferred from whole mitogenome sequences. Biol J Linn Soc. 98:923–936.