

Characterization of mitochondrial mRNAs in codfish reveals unique features compared to mammals

Dag H. Coucheron · Marianne Nymark · Ragna Breines · Bård Ove Karlsen · Morten Andreassen · Tor Erik Jørgensen · Truls Moum · Steinar D. Johansen

Received: 19 January 2011 / Revised: 27 March 2011 / Accepted: 28 March 2011 / Published online: 12 April 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Expression and processing of mitochondrial gene transcripts are fundamental to mitochondrial function, but information from early vertebrates like teleost fishes is essentially lacking. We have analyzed mitogenome sequences of ten codfishes (family Gadidae), and provide complete sequences from three new species (Saithe, Pollack and Blue whiting). Characterization of the mitochondrial mRNAs in Saithe and Atlantic cod identified a set of ten poly(A) transcripts, and six UAA stop codons are generated by posttranscriptional polyadenylation. Structural assessment of poly(A) sites is consistent with an RNaseP cleavage activity 5' of tRNA acceptor-like stems. COI, ND5 and ND6 mRNAs were found to harbor 3' UTRs with antisense potential extending into neighboring gene regions. While the 3' UTR of COI mRNA is complementary to the tRNA^{Ser (UCN)} and highly similar to that detected

in human mitochondria, the ND5 and ND6 3' UTRs appear more heterogenic. Deep sequencing confirms expression of all mitochondrial mRNAs and rRNAs, and provides information about the precise 5' ends in mature transcripts. Our study supports an overall evolutionary conservation in mitochondrial RNA processing events among vertebrates, but reveals some unique 5' and 3' end characteristics in codfish mRNAs with implications to antisense regulation of gene expression.

Keywords Antisense regulation · Deep sequencing · Gadidae · Mitochondrial RNA · Mitogenome · Polyadenylation

Introduction

Vertebrate mitochondrial genomes (mitogenomes) are usually 16–17 kb in size and relatively uniform in gene content and structure. The circular and multicopy mitogenomes encode only 37 of the about 1,500 gene products essential for mitochondrial structure and function (Wallace 2007). More than 1,400 complete vertebrate mitogenome sequences are available in public databases (<http://www.ncbi.nlm.nih.gov/genome>) all containing the same set of genes representing the small- and large-subunit ribosomal RNA (mtSSU and mtLSU rRNAs), 22 transfer RNAs (tRNAs) and 13 respiratory chain proteins. These include seven NADH dehydrogenase subunits (ND1–6, ND4L), three cytochrome c oxidase subunits (COI–III), two ATP synthase subunits (A6 and A8), and the cytochrome b (CytB). The mitogenome organization in general is highly conserved from fishes to mammals; only minor differences are noted in birds, reptiles, marsupials, jawless fishes, and a few other lineages that include some fish genera (e.g. Inoue

Communicated by L. Tomaska.

The complete mitogenome sequences from Saithe (*Pollachius virens*), Pollack (*P. pollachius*), and Blue whiting (*Micromesistius poutassou*) are available from the EMBL Database under the accession numbers FR751399, FR751400, and FR751401, respectively.

Electronic supplementary material The online version of this article (doi:10.1007/s00294-011-0338-2) contains supplementary material, which is available to authorized users.

D. H. Coucheron · M. Nymark · R. Breines · B. O. Karlsen · M. Andreassen · S. D. Johansen (✉)
RNA and Transcriptomics group, Department of Medical Biology, Faculty of Health Sciences, University of Tromsø, MH-building Breivika, 9037 Tromsø, Norway
e-mail: Steinar.Johansen@uit.no

B. O. Karlsen · T. E. Jørgensen · T. Moum · S. D. Johansen
Marine Genomics group, Faculty of Biosciences and Aquaculture, University of Nordland, Bodø, Norway

et al. 2003; Satoh et al. 2006; Breines et al. 2008). This conserved structural pattern strongly indicates a common pathway of gene expression.

Transcription of vertebrate mitochondrial genes is well studied in human cells and tissues. Both the heavy (H) and light (L) strands are transcribed almost completely and symmetrically from promoters located within or close to the mitochondrial control region (CR). Two distinct promoters (H₁ and H₂) are responsible for the H-strand specific transcription, and one single promoter for the L-strand (Montoya et al. 1982, 1983; Fernandez-Silva et al. 2003; Falkenberg et al. 2007; Temperley et al. 2010b). The H₂ promoter is located within the tRNA^{Phe} gene and generates a large polycistronic RNA that is further processed into 10 H-strand specific mRNAs, 2 rRNAs, and 13 tRNAs. In addition, the rRNA genes and two tRNA genes are transcribed from the H₁ promoter, including the tRNA^{Phe} which is not part of the H₂ transcript. The H₁-specific transcript is short and abundant, and terminates within the tRNA^{Leu (UUR)} gene downstream of the mtLSU rRNA gene at the mTERF binding site (Valverde et al. 1994; Hyvarinen et al. 2007). The long polycistronic L-strand specific transcript is initiated at the L-strand promoter within the CR and gives rise to the ND6 mRNA, eight tRNAs, and a small CR-specific polyadenylated RNA species (7S RNA) with unknown function.

All mature mitochondrial mRNAs are monocistronic except for two bicistronic mRNAs corresponding to the ND4/ND4L subunits and the A8/A6 subunits. Human mitochondrial mRNAs typically lack 5' and 3' untranslated regions (UTRs), but some notable exceptions have been reported. The COI mRNA has a 3' UTR of 72 nt complementary to the complete tRNA^{Ser (UCN)} sequence, COII has a 15–24 nt 3' UTR that depends on cell type, and ND5 appears to have an extensive 3' UTR complementary to the ND6 mRNA (Slomovic et al. 2005; Temperley et al. 2010b). Similarly, ND6 mRNA is highly heterogeneous at the 3' end with long 3' UTRs complementary to the ND5 mRNA (Slomovic et al. 2005). Finally, most human mitochondrial mRNAs contain short tails of approximately 40–45 adenosines added post-transcriptionally at, or immediately after, the stop codon. Seven UAA stop codons in human mitochondria are generated by polyadenylation (Nagaike et al. 2005; Slomovic et al. 2005; Nagaike et al. 2008; Borowski et al. 2010; Temperley et al. 2010b). Interestingly, there is cell-specific heterogeneity in mRNA polyadenylation in human mitochondria and notable examples are found in the ND5 and ND6 mRNAs (see Temperley et al. 2010b). Mitochondrial mRNAs have been investigated in a few additional mammals (e.g. Tullo et al. 1994; Bai et al. 2000; Piruat and Lopez-Barneo 2005) and these results corroborate the main findings in humans. Detailed knowledge on mitochondrial transcription

products of non-mammalian vertebrate species is essentially lacking. Characterization based on the Northern hybridization of transcription products in rainbow trout has been performed in liver mitochondria (Zardoya et al. 1995), and the overall pattern supports that of mammals. Similarly, we reported that mtSSU and mtLSU rRNAs in Atlantic cod are oligoadenylated at their 3' ends (Bakke and Johansen 2002), a finding similar to that in human mitochondrial rRNA (Dubin et al. 1982; Temperley et al. 2010b).

The family Gadidae comprises several commercially and ecologically important species that are mainly confined to the North Atlantic Ocean. Complete Gadidae mitogenome sequences have been reported previously from seven species (Table 1). Here we expand on the mitogenomic analyses of gadid fishes by reporting the complete mitogenome sequences of three additional gadids (Saithe *Pol-lachius virens*; Pollack *P. pollachius*; Blue whiting *Micromesistius poutassou*) and characterizing the mitochondrial mRNAs in two representative gadids (Saithe; Atlantic cod) by RT-PCR sequencing and 454 pyrosequencing.

Materials and methods

Gadid fish samples, nucleic acid extraction, and PCR amplification

The Saithe, Pollack, and Blue whiting specimens were collected from the coastal area of mid and northern Norway. DNA for complete mitogenome sequencing was extracted as described from frozen muscle tissue using the mtDNA Extractor CT Kit from Wako (Ursvik et al. 2007). Total RNAs from Saithe and Atlantic cod were isolated from fresh muscle tissue essentially as described previously (Bakke and Johansen 2002) and applied in reverse transcriptase (RT) PCR sequencing and pyrosequencing library constructions. The PCR and sequencing primers were designed from the published *G. morhua* mtDNA sequence (Johansen and Bakke 1996) as well as other available gadid mtDNA sequences (Table 1) and used to amplify the mitogenomes in to 1–4 kb fragments essentially as described in Breines et al. (2008).

DNA sequencing

Sanger sequencing of mitogenomes was performed directly on PCR products on both strands as previously described (Breines et al. 2008) using the BigDye kit (Applied Biosystems), with the same primers as in the PCR and internal primers. Roche 454 pyrosequencing was performed as a service given by Eurofins MWG Operon (Germany). Here,

Table 1 Complete mitogenome sequences in gadid species (family Gadidae)

Species	Ref.	Accession number
<i>Arctogadus glacialis</i> (Arctic cod)	Breines et al. (2008)	AM919429
<i>Boregadus saida</i> (Polar cod)	Breines et al. (2008)	AM919428
^a <i>Gadus ogac</i> (Greenland cod)	Coulson et al. (2006)	DQ356940
^a <i>Gadus macrocephalus</i> (Pacific cod)	Coulson et al. (2006)	DQ356937
<i>Gadus morhua</i> (Atlantic cod)	Johansen and Bakke (1996); Ursvik et al. (2007)	X99772; AM489716
<i>Melanogrammus aeglefinus</i> (Haddock)	Roques et al. (2006); Ursvik et al. (2007)	DQ020497; AM489717
<i>Merlangius merlangus</i> (Whiting)	Roques et al. (2006)	DQ020496
^a <i>Microgadus proximus</i> (Pacific tomcod)	Coulson et al. (2006)	DQ356944
<i>Micromesistius poutassou</i> (Blue whiting)	This work	FR751401
<i>Pollachius pollachius</i> (Pollack)	This work	FR751400
<i>Pollachius virens</i> (Saithe)	This work	FR751399
<i>Theragra chalcogramma</i> (Walleye pollock)	Yanagimoto et al. (2004)	AB094061; AB182300–08
<i>Theragra finnmarchia</i> (Norwegian pollock)	Ursvik et al. (2007)	AM489718; AM489719

^a Mitogenome not completely sequenced. The control region, TP-spacer and some tRNA gene sequences are missing

polyA-enriched normalized cDNA libraries (based on total cellular RNA) were generated by random hexamer first strand synthesis. About 160,000 reads were obtained from equal amounts of pooled mRNA isolated from each of several tissue types of adult Atlantic cod (liver, muscle, heart, spleen and brain). Similarly 160,000 reads were obtained from pooled mRNAs for 16 developmental stages from zygote to larva (Johansen et al. 2009, 2011). In addition, 1.2 million reads were obtained from a liver-specific cDNA library generated by pooling equal amounts of total RNA from ten individuals of Atlantic cod.

Data analyses

Computer analyses of DNA sequences were performed using software from DNASTAR Inc. For the phylogenetic analyses two sets of multiple alignments of nucleotide sequences derived from the complete mitochondrial genomes of ten Gadidae species were generated using Clustal X v/1.81 (Thompson et al. 1997) with manual refinements. One alignment was based on 16,636 nt positions of the complete mitochondrial genome, except for the highly variable ETAS (extended termination associated sequence) domain within CR. The second alignment was built from the complete set of codons (except stop codons) creating a concatenated sequence of 11,406 nt positions (3,802 codons) corresponding to the 13 protein genes. The tree-building methods of minimum evolution (ME) and maximum parsimony (MP) in MEGA version 4.0 (Tamura et al. 2007), as well as maximum likelihood (ML) in PAUP* version 4.0b10 (Swofford 2002) were used to reconstruct molecular phylogenies. ME trees were built using different nucleotide substitution models. MP trees were reconstructed using the Max-mini Branch-&-bound search

option. ML trees were built from sequence evolution models generated by the computer program WinModeltest version 4b (Posada and Crandall 1998). The topologies of the ME, MP and ML trees were evaluated by bootstrap analyses (2,000 replications). The membrane topologies of the 13 mitochondrial proteins were predicted using different methods available as web servers, including ConPrepII, Phobius, and TOPCONS (Arai et al. 2004; Käll et al. 2007; Bernsel et al. 2009). The total 1.52 million 454 pyrosequencing reads were then formatted and fed into a local blast database and subsequently mined for contigs corresponding to the complete protein gene regions of the Atlantic cod mitogenome (<ftp://ftp.ncbi.nih.gov/blast/executables/release/LATEST/>). The pyrosequencing reads mapped to the mitogenome were visualized by the Tablet graphical viewer (Milne et al. 2009).

Results and discussion

Complete sequencing of gadid mitogenomes and phylogenetic relationships

The complete mitogenome sequences of three new gadid species (family Gadidae, order Gadiformes) are provided by this study, and include Saithe (*Pollachius virens*; 16,556 bp), Pollack (*P. pollachius*; 16,539 bp), and Blue whiting (*Micromesistius poutassou*; 16,573 bp). General structural features of the mitogenomes, including size, gene content and organization were similar to those of other published gadids and vertebrates in general (Johansen and Bakke 1996; Yanagimoto et al. 2004; Roques et al. 2006; Ursvik et al. 2007; Falkenberg et al. 2007; Breines et al. 2008). The number of complete gadid mitogenome

sequences now available is approximately 25, and represents 10 species. Extensive partial sequences have been reported for three more species (Table 1) (Coulson et al. 2006). Some gadid mitogenomes contain heteroplasmic tandem repeat (HTR) sequences within the main control region (Atlantic cod), at the light-strand origin of replication (Polar cod *Boreogadus saida*), or within the spacer region between the tRNA^{Thr} and tRNA^{Pro} genes (Arctic cod *Arctogadus glacialis*) (Figure S1A) (Johansen et al. 1990; Árnason and Rand 1992; Breines et al. 2008; Pálsson et al. 2008), but no mitochondrial HTRs were noted either in the Saithe, Pollock, or Blue whiting. The derived mitochondrial protein sequences (13 membrane protein subunits) from the 10 gadid species were aligned and assessed for amino acid substitutions. Among the 184 variable amino acid positions detected, 20 positions were represented by three or more variant amino acids and the majority of substitutions were only present in Blue whiting (Figure S2). About 50% of the mitochondrial amino acid residues were found to be present in trans-membrane (TM) regions, and substitutions were equally distributed among TM and non-TM regions. The complex IV subunits (COI–III) were found to be approximately 10 times more conserved compared to most other mtDNA-encoded proteins.

To investigate the taxonomic relationship among the gadids we performed high-resolution phylogeny based on mitochondrial concatenated protein-gene (11,406 positions corresponding to 3,802 amino acid positions) and complete mitogenome sequences (16,636 positions) from the ten species. High-resolution phylogenetic analyses resolved all genera with strong statistical support and all trees generated resulted in essentially identical tree topologies. A representative ML tree is presented in Figure S1B, and supports the overall relationships among gadids reported in recent works (Bakke and Johansen 2005; Coulson et al. 2006; Teletchea et al. 2006; Breines et al. 2008; Roa-Varon and Orti 2009). The phylogenetic analysis strongly supports sister-taxa affiliation among the gadids. The *Micromesistius* genus appears as the most diverged taxon, and *Theragra gadus*, *Arctogadus/Boreogadus*, *Merlangius melanogrammus*, and the two *Pollachius* species form distinct and more recent clades within the Gadidae tree (Figure S1B).

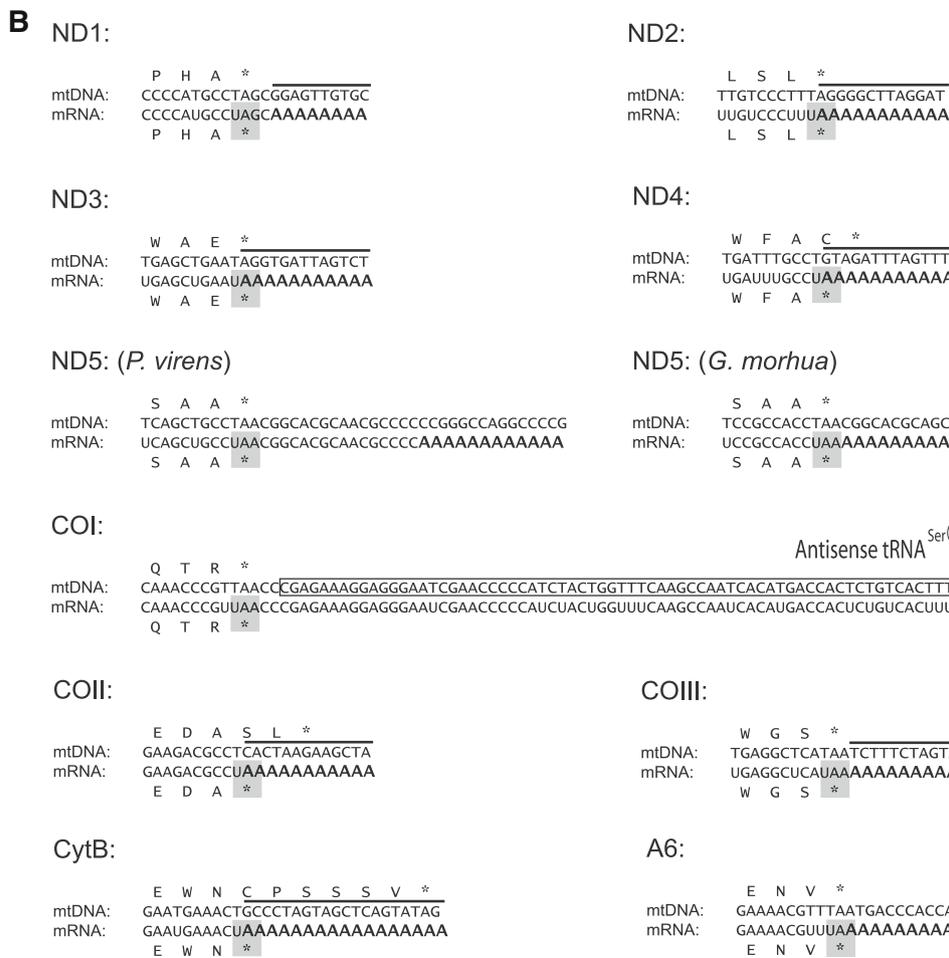
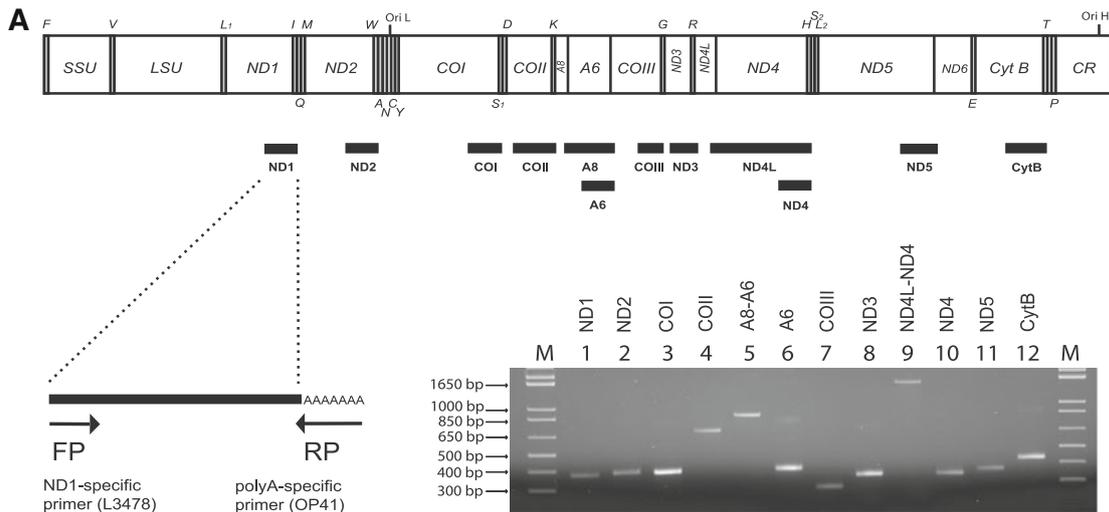
Mapping of poly(A) sites at the 3' end of mt-mRNAs based on RT-PCR sequencing

To map and characterize poly(A) sites in mitochondrial transcripts of gadids, an RT-PCR sequencing approach was applied to Saithe and Atlantic cod. Here, total RNA was subjected to reverse transcription by a poly(A)-specific reverse primer and subsequently amplified by the addition of a gene-specific primer (Fig. 1a). Amplified products

Fig. 1 Mapping of poly(A) sites at the 3' ends of H-strand specific mitochondrial mRNAs in Saithe (*P. virens*) and Atlantic cod (*G. morhua*). **a** Experimental strategy indicated on the linear map of the circular mtDNA (see legends to Figure S1 for abbreviations). Amplified and analyzed regions of mRNAs are indicated as *black bars* below the map. The mRNA species were amplified by RT-PCR using a poly(A)-specific reverse primer (RP) and an mRNA-specific forward primer (FP). Amplification gave one single major cDNA fragment for each RNA species as evaluated by agarose gel electrophoresis (*below right*). Note that all mRNAs, except two, are monocistronic. The A8/A6 mRNA and the ND4L/ND4 mRNA are both bicistronic. **b** Summary of poly(A) site sequence mapping, including a comparison of corresponding mRNA and mtDNA sequences. Poly(A) tails added post-transcriptionally are indicated, and *asterisk* refers to stop codons. Note that in several cases the stop codon inferred from the mtDNA sequence does not match the experimental determined stop codons in mRNA. *Grey lines* above mtDNA sequences indicate neighboring tRNA genes. The antisense sequence of tRNA^{Ser} (UCN) is shown as a 3' UTR of the COI mRNA. The sequences shown are from Saithe, but all results except for the ND5 mRNA are identical in Saithe and Atlantic cod

were detected from all the H-strand specific mRNAs, but not the L-strand specific ND6 mRNA. The latter observation is consistent with the lack of a stable poly(A) tail in mammalian mitochondrial ND6 mRNA (Slomovic et al. 2005; Temperley et al. 2010b). Furthermore, similar to mammals, the gadid fishes generated eight monocistronic transcripts and two bicistronic H-strand specific transcripts corresponding to the A6/A8 and ND4L/ND4. Bicistronic transcripts gave identical poly(A) sites from both gene-specific primers (e.g. from A8 and A6, and from ND4L and ND4).

Several interesting findings could be derived from the sequences summarized in Fig. 1b. First, the poly(A) site in most cases corresponded exactly to the 5' end of the downstream tRNA, consistent with the punctuation model of mammalian mtDNA (Ojala et al. 1981). Second, all H-strand specific stop codons (except that of ND1 mRNA) were UAA, where at least six are generated by posttranscriptional polyadenylation. Third, three of the messengers indicated shorter protein products compared to that derived from the corresponding DNA sequence. ND4, COII, and CytB appeared to be 1, 2, and 6 amino acids shorter, respectively. Finally, two of the transcripts (ND5 and COI mRNAs) contained extended 3' UTRs. The 3' UTR of COI was complementary to the complete tRNA^{Ser} (UCN), and corresponded to 76 nt in both the Saithe and Atlantic cod. Heterogeneity in poly(A) sites was detected when comparing the ND5 mRNAs of Saithe and Atlantic cod. While polyadenylation occurs at the UAA stop codon in Atlantic cod, a 16 nt 3' UTR complementary to the ND6 coding region was present in Saithe (Fig. 1b). Heterogeneity at the 3' end of the ND5 mRNA has been reported in human mitochondria (Temperley et al. 2010b). Different from that reported in human mitochondria, we did not find any 3' UTR in the COII mRNA of Atlantic cod and Saithe.



All poly(A) sites mapped exactly at the 5' end of a downstream RNA helical structure (Fig. 2a). With the notable exceptions of the A6 and ND5 mRNAs, this helix corresponds to the acceptor stem of a mitochondrial tRNA. The A6 and ND5 genes, however, have no tRNA

sequences immediately downstream, but have protein-coding gene regions (COIII and ND6, respectively). Interestingly, the A6 poly(A) site corresponds exactly to the AUG start codon of the COIII messenger at a paired structure similar to that of a tRNA acceptor stem

(Fig. 2a). Similarly, the ND5 mRNA poly(A) sites mapped at paired structures located complementary to the ND6 mRNA. The ribosomal RNAs are oligoadenylated at their 3' ends according to downstream tRNA acceptor structures (Fig. 2b) (Bakke and Johansen 2002). In summary we infer that poly(A) site cleavage of mRNA and 5' end processing of tRNA are due to one single cleavage event. Here, mitochondrial RNaseP is an obvious candidate because of its tRNA acceptor stem specificity and the fact that it leaves a free 3'OH end for poly(A) polymerases priming (Xiao et al. 2002; Nagaike et al. 2008; Esakova and Krasilnikov 2010).

Assessment of abundance and 5' ends of mRNAs by 454 pyrosequencing

Deep sequencing was included as an additional strategy for analyzing mitochondrial transcripts. About 1.52 million reads obtained by 454 pyrosequencing representing the Atlantic cod transcriptome were assessed for mitochondrial transcripts. These sequences were generated from temporal and spatial RNA samplings during developmental and adult stages, respectively (see “Materials and methods”). A total of 4,698 reads were unambiguously identified as mitochondrial transcripts, and among these 2,328 reads

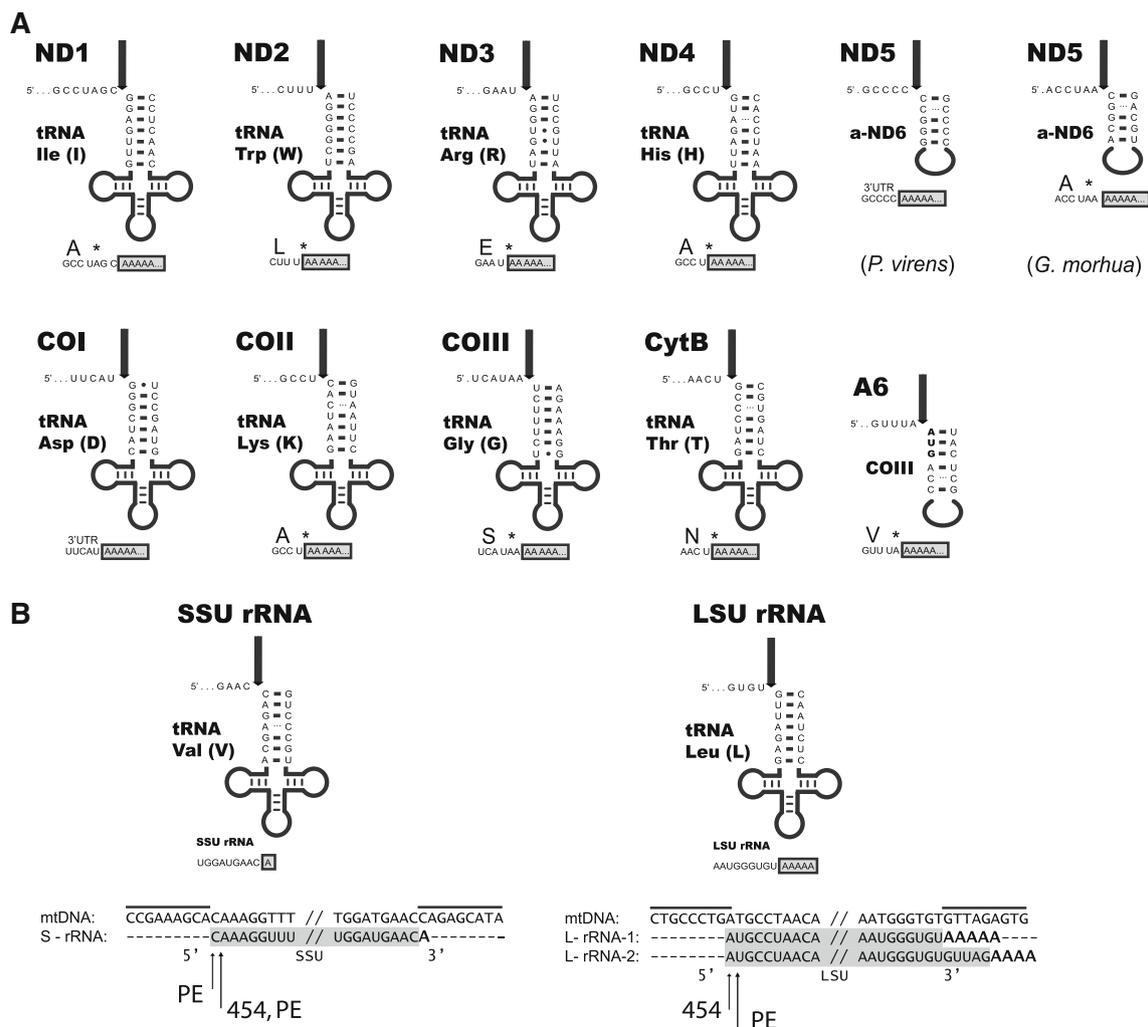


Fig. 2 Correlation between polyA sites and downstream tRNA-like structures. **a** Schematic presentation of the 3' end polyA sites in all H-strand specific mRNAs. The polyA sites map exactly 5' of the acceptor stem of tRNAs (ND1, ND2, ND3, ND4, COI, COII, COIII and CytB mRNAs) or a stem structure apparently similar to that of a tRNA (ND5 and A6 mRNAs). The structural features from Saithe are shown, but all results except for the ND5 mRNA are identical in Saithe and Atlantic cod. See legend to Fig. 2 for more details. **b** Oligoadenylation sites at the Atlantic cod rRNA 3' ends follow

similar structural rules as the mRNAs. Note that two different adenylation sites are present at the LSU rRNA 3' end, but only the L-rRNA-2 transcript is processed according to the tRNA acceptor stem structure. The L-rRNA-1 is part of a short primary mitochondrial transcript initiated from the H₁ promoter and terminated within the tRNA^{Leu} gene by the mTERF protein. The 5' ends of mt-SSU and mt-LSU rRNAs were determined by primer extension (PE) analysis and by 454 pyrosequencing of RNA (454)

mapped to protein-coding regions (Fig. 3a, b). Here, the L-strand specific ND6 mRNA and all ten H-strand specific mRNAs were represented by multiple reads each. When normalizing the read numbers against actual gene size (reads/kb) we note that most messengers were represented at similar level, but some transcripts (e.g. A8/A6 and ND6) appeared at slightly lower abundance than others (e.g. ND2, ND5 and COI). Interestingly, we observed multiple reads of an L-strand specific non-coding transcript (LncRNA) covering the CSB (conserved sequence box) domain and parts of the CC (central conserved) domain of the CR (see Lee et al. 1995 for CR domain definition). LncRNA is apparently homologous to the 7S RNA previously detected in human mitochondria (Fig. 3a) (Gaines and Attardi 1984). We also observed multiple pyrosequencing reads corresponding to long

extended 3' UTR sequences of the L-strand specific ND6 mRNA.

A closer inspection of the 2,328 mitochondrial-specific reads identified a number of sequences that correspond to the 5' and 3' ends of the mRNAs. The 3' polyadenylation sites indicated by deep sequencing are in agreement with those of the RT-PCR sequencing approach presented above. The 5' end maps are summarized in Fig. 3c and indicate that most mRNAs (7 of 11) contain short 5' UTRs of one or two nucleotides upstream of the AUG start codon. This feature is different from that of human mitochondria where 8 of 11 mRNAs lack 5' UTR nucleotides (Temperley et al. 2010b). Only the 5' ends of ND5 and ND6 mRNAs are identical in cod and human, and start exactly at the AUG initiation codon. An interesting example was noted at the COII mRNA 5' end. Here, the major end contained a

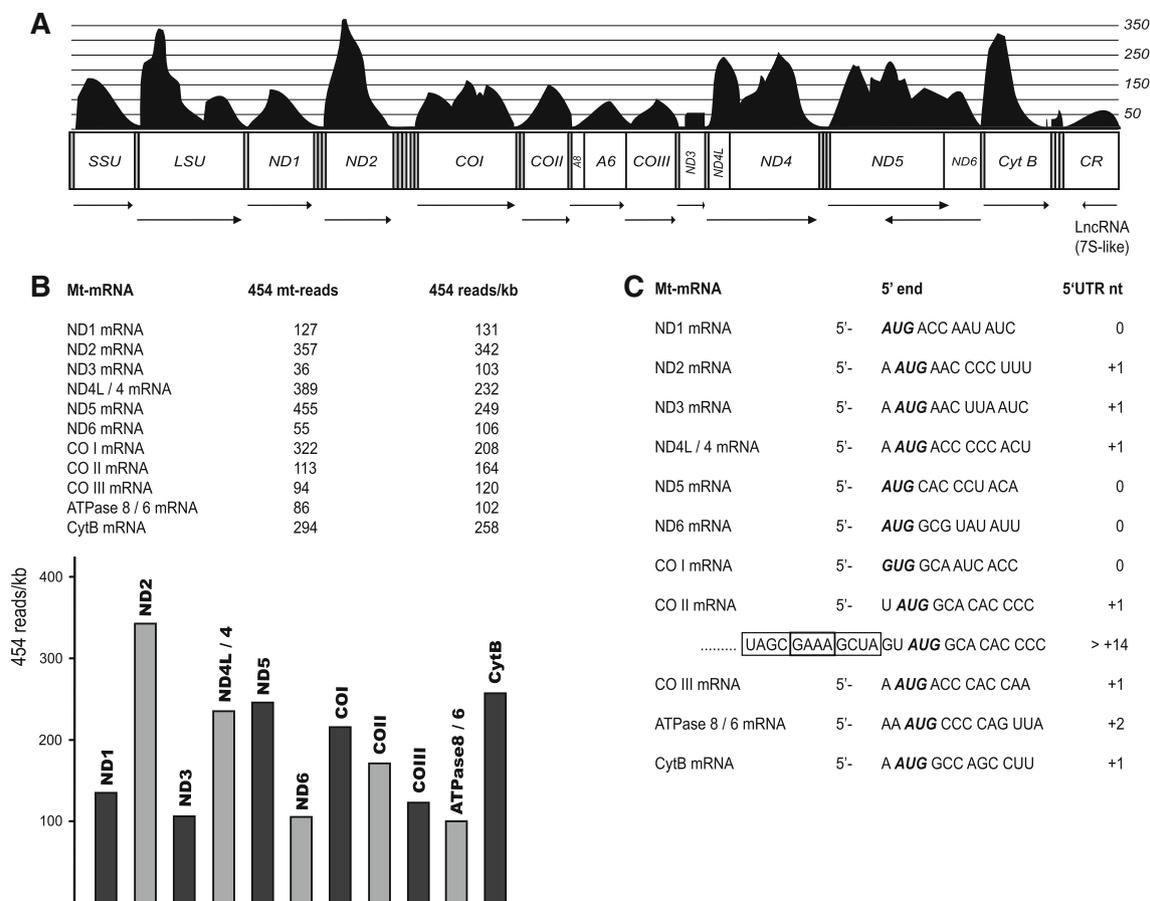


Fig. 3 Mining mitochondrial mRNA transcripts from 454 pyrosequencing libraries in Atlantic cod. Libraries were based on the poly(A) fraction of total cellular RNA isolated from adult tissues and various development stages. **a** Mapping of 454 reads from protein-coding regions, rRNA coding regions, and CR to the mitogenome. Coverage per nucleotide position is presented by the *histograms* above the mitogenome organization map. *Arrows* below the organization map pointing right and left represent the mature H-strand specific and L-strand specific transcripts, respectively. The LncRNA

in CR corresponds to the human 7S RNA. **b** The 454 read numbers of each mRNA are normalized according to size of gene coding regions (454 reads/kb). The estimated normalized read numbers vary between approximately 100 and 300 reads/kb (schematically presented below). **c** Mapping the 5' ends of mitochondrial mRNAs by pyrosequencing. The 5' ends corresponding to the majority of reads for each gene are shown. Note that some transcripts of the COII mRNA include upstream sequence corresponding to a GAAA tetraloop hairpin (*boxed*)

one nucleotide 5' UTR not present in human mitochondria, as well as minor 5' ends of at least 14 nt that include upstream sequences covering a highly conserved noncoding spacer between the tRNA^{ASP} and COII gene regions. This spacer corresponds to a hairpin structure present in all vertebrates, and in all codfish species investigated the hairpin is capped by a GAAA tetraloop (Fig. 3c). Tetraloops of the GNRA family (N is any base, R is a purine) have unique RNA structural organization and frequently participates in long-range tertiary RNA:RNA interactions (Abramovitz and Pyle 1997).

Pyrosequencing identified 46 and 120 reads corresponding to the 5' ends of mtSSU rRNA and mtLSU rRNA, respectively. Based on primer extension (PE) experiments we reported previously two adjacent 5' ends of mtSSU rRNA and one single 5' end of mtLSU rRNA (Bakke and Johansen 2002). Analysis of the 454 reads revealed a single 5' end in mtSSU rRNA (46/46 reads) corresponding to one of the PE sites. Similarly, 119 of the 120 mtLSU rRNA reads mapped one position downstream that of the PE site (Fig. 2b).

Concluding remarks

The mitochondrial gene content and organization is shared by most vertebrates from fishes to mammals, and here we report that the mitochondrial mRNA processing patterns in codfishes are in overall agreement with that of humans. However, some unique characteristics are noted at the 5' and 3' ends of fish mitochondrial mRNA transcripts. We identified three mRNA 3' UTRs of significant size in codfish with antisense potential to other coding sequences in the mitogenome (Fig. 4). The COI 3' UTR corresponds to the full complementary sequence of tRNA^{Ser(UCN)}. This feature was unambiguously detected in both Saithe and Atlantic cod, and is similar to the 3' UTR previously characterized as the single form of COI mRNA in human mitochondria (Slomovic et al. 2005). This molecular feature appears highly conserved, and is probably present in all vertebrate mitochondria. No function has yet been assigned to the COI 3' UTR, but one possibility is in regulation of processing events due to antisense stabilization between H-strand and L-strand gene transcripts. Interestingly, a mutation associated with hearing loss in humans has been identified in the tRNA^{Ser(UCN)} precursor that significantly affects RNA processing and stability of this particular tRNA as well as a marked reduction of the ND6 mRNA level (Guan et al. 1998).

The last two examples of 3' UTRs include both the ND5 mRNA and the ND6 mRNA. The 3' end in ND6 of mammals is several hundred nucleotides and complementary to the coding part of ND5 mRNA (Tullo et al. 1994).

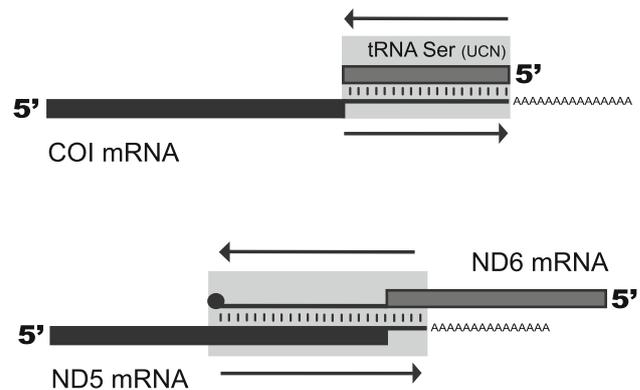


Fig. 4 Schematic presentations of antisense regions between the H-strands and L-strand specific mature RNAs. The COI mRNA contains a 76 nt 3' UTR complementary to the L-strand specific tRNA^{Ser(UCN)}. The 3' UTR of ND6 mRNA is heterogenic in size among mammals and is complementary to hundreds of nucleotides of the H-strand specific ND5 mRNA. ND6 mRNA contains no conserved poly(A) site

Our pyrosequencing experiment in Atlantic cod reveals a similar feature at the ND5 mRNA 3' end. We noted a short 3' UTR of ND5 mRNA complementary to the coding part of ND6 mRNA in Saithe (Figs. 1b, 4), resulting in a potential base pairing between the H-strand specific ND5 mRNA and the L-strand specific ND6 mRNA. Studies in mammalian mitochondria have revealed deviant features of the ND5 and ND6 mRNAs compared to the other mitochondrial mRNAs. The ND5 gene expression is the only mitochondrial protein-coding gene that is tightly regulated, and the rate of ND5 synthesis appears to control the overall rate of respiration (Bai et al. 2000; Chomyn 2001). Interestingly, this regulation appears due to differential ND5 mRNA stabilization, and mRNA levels have been reported to be upregulated by nutrients or cholera pathogenesis, and downregulated during hypoxia (Everts and Berdanier 2002; Piruat and Lopez-Barneo 2005; Sarkar et al. 2005). The ND6 mRNA is the only protein-coding mRNA generated from the L-strand specific primary transcript. This messenger has a very long 3' UTR (ca 600 nt in rat mitochondria and probably of similar size in Atlantic cod), performs stop codon frameshifting, and contains no defined 3' end with regular poly(A) tail (Tullo et al. 1994; Nagaïke et al. 2008; Temperley et al. 2010a, b). Furthermore, the ND6 mRNA is closely associated with the mitochondrial DNA, a feature different from the H-strand specific mRNAs (Ozawa et al. 2007). The fact that both the ND5 and ND6 mRNAs have unusual short poly(A) tails compared to other mitochondrial mRNAs (Temperley et al. 2010b) suggests a role in antisense stabilization of the messengers. Our study corroborates findings in mammals that antisense RNA transcripts appear to be involved in mitochondrial function (Sbisa et al. 1992; Tullo et al. 1994; Sbisa et al. 1997; Slomovic et al. 2005; Lung et al. 2006).

Acknowledgments We thank Anita Ursvik, Tomasz Furmanek, Åse Emblem, and Jarle T. Nordeide for discussions, and Eurofins MWG Operon for pyrosequencing services. This work was supported by grants from the MABIT (Marine Biotechnology in Northern Norway) program, the Norwegian Research Council by the GenoFish Consortium, the University of Nordland, and the University of Tromsø.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Abramovitz DL, Pyle AM (1997) Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. *J Mol Biol* 266:493–506
- Arai M, Mitsuke H, Ikeda M, Xia J-X, Kikuchi T, Satake M, Shimizu T (2004) ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic Acids Res* 32:W390–W393
- Árnason E, Rand DM (1992) Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. *Genetics* 132:211–220
- Bai Y, Shakeley RM, Attardi G (2000) Tight control of respiration by NADH dehydrogenase ND5 subunit gene expression in mouse mitochondria. *Mol Cell Biol* 20:805–815
- Bakke I, Johansen S (2002) Characterization of mitochondrial ribosomal RNA genes in gadiformes: sequence variations, secondary structural features, and phylogenetic implications. *Mol Phylogenet Evol* 25:87–100
- Bakke I, Johansen SD (2005) Molecular phylogenetics of gadidae and related gadiformes based on mitochondrial DNA sequences. *Mar Biotechnol* 7:61–69
- Bernsel A, Viklund H, Hennerdal A, Elofsson A (2009) TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res* 37:W465–W468
- Borowski LS, Szczesny RJ, Brzezniak LK, Stepień PP (2010) RNA turnover in human mitochondria: more questions than answers? *Biochim Biophys Acta* 1797:1066–1070
- Breines R, Ursvik A, Nymark M, Johansen SD, Coucheron DH (2008) Complete mitochondrial genome sequences of the Arctic Ocean codfishes *Arctogadus glacialis* and *Boreogadus saida* reveal oriL and tRNA gene duplications. *Polar Biol* 31:1245–1252
- Chomyn A (2001) Mitochondrial genetic control of assembly and function of complex I in mammalian cells. *J Bioenerg Biomembr* 33:251–257
- Coulson MW, Marshall HD, Pepin P, Carr SM (2006) Mitochondrial genomics of gadine fishes: implications for taxonomy and biogeographic origins from whole-genome data sets. *Genome* 49:1115–1130
- Dubin DT, Montoya J, Timko KD, Attardi G (1982) Sequence analysis and precise mapping of the 3' ends of HeLa cell mitochondrial ribosomal RNAs. *J Mol Biol* 157:1–19
- Esakova O, Krasilnikov AS (2010) Of proteins and RNA: the RNase P/MRP family. *RNA* 16:1725–1747
- Everts HB, Berdanier CD (2002) Regulation of mitochondrial gene expression by retinoids. *IUBMB Life* 54:45–49
- Falkenberg M, Larsson NG, Gustafsson CM (2007) DNA replication and transcription in mammalian mitochondria. *Annu Rev Biochem* 76:679–699
- Fernandez-Silva P, Enriques JA, Montoya J (2003) Replication and transcription of mammalian mitochondrial DNA. *Exp Physiol* 88.1:41–56
- Gaines G, Attardi G (1984) Highly efficient RNA-synthesizing system that uses isolated human mitochondria: new initiation events and in vivo-like processing patterns. *Mol Cell Biol* 4:1605–1617
- Guan MX, Enriquez JA, Fischel-Ghodsian N, Puranam RS, Lin CP, Maw MA, Attardi G (1998) The deafness-associated mitochondrial DNA mutation at position 7445, which affects tRNA^{Ser(UCN)} precursor processing, has long-range effects on NADH dehydrogenase subunit ND6 gene expression. *Mol Cell Biol* 18:5868–5879
- Hyvarinen AK, Pohjoismaki JL, Reyes A, Wanrooij S, Yasukawa T, Karhunen PJ, Spelbrink JN, Holt IJ, Jacobs HT (2007) The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. *Nucleic Acids Res* 35:6458–6474
- Inoue JG, Miya M, Tsukamoto K, Nishida M (2003) Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the “ancient fish”. *Mol Phylogenet Evol* 26:110–120
- Johansen S, Bakke I (1996) The complete mitochondrial DNA sequence of Atlantic cod (*Gadus morhua*): relevance to taxonomic studies among codfishes. *Mol Mar Biol Biotechnol* 5:203–214
- Johansen S, Guddal PH, Johansen T (1990) Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua*. *Nucleic Acids Res* 18:411–419
- Johansen SD, Coucheron DH, Andreassen M, Karlsen BO, Furmanek T, Jørgensen TE, Emblem Å, Breines R, Nordeide JT, Moum T, Nederbragt AJ, Stenseth NC, Jakobsen KS (2009) Large-scale sequence analyses of Atlantic cod. *N Biotechnol* 25:263–271
- Johansen SD, Karlsen BO, Furmanek T, Andreassen M, Jørgensen TE, Bizuayehu TT, Breines R, Emblem Å, Kettunen P, Luukko K, Edvardsen RB, Nordeide JT, Coucheron DH, Moum T (2011) RNA deep sequencing of the Atlantic cod transcriptome. *Comp Biochem Physiol Part D* 6:18–22
- Käll L, Krogh A, Sonnhammer ELL (2007) Advantages of combined transmembrane topology—the Phobius web server. *Nucleic Acids Res* 35:W429–W432
- Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *J Mol Evol* 41:54–66
- Lung B, Zemann A, Madej MJ, Schuelke M, Techritz S, Ruf S, Bock R, Huttenhofer A (2006) Identification of small non-coding RNAs from mitochondria and chloroplasts. *Nucleic Acids Res* 34:3842–3852
- Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, Marshall D (2009) Tablet—next generation sequence assembly visualization. *Bioinformatics* 26:401–402
- Montoya J, Christianson T, Levens D, Rabinowitz M, Attardi G (1982) Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proc Natl Acad Sci USA* 79:715–7199
- Montoya J, Gaines GL, Attardi G (1983) The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell* 34:151–159
- Nagaike T, Suzuki T, Katoh T, Ueda T (2005) Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase. *J Biol Chem* 280:19721–19727
- Nagaike T, Suzuki T, Ueda T (2008) Polyadenylation in mammalian mitochondria: insight from recent studies. *Biochem Biophys Acta* 1779:266–269

- Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290:470–474
- Ozawa T, Natori Y, Sato M, Umezawa Y (2007) Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat Methods* 4:413–419
- Pálsson S, Paulsen J, Árnason E (2008) Rapid evolution of the intergenic T-P spacer in the mtDNA of Arctic cod *Arctogadus glacialis*. *Mar Biotechnol* 10:270–277
- Piruat JI, Lopez-Barneo J (2005) Oxygen tension regulates mitochondrial DNA-encoded complex I gene expression. *J Biol Chem* 280:42676–42684
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818
- Roa-Varon A, Orti G (2009) Phylogenetic relationships among families of Gadiformes (Teleostei, Paracanthopterygii) based on nuclear and mitochondrial data. *Mol Phylogenet Evol* 52:688–704
- Roques S, Fox CJ, Villasana MI, Rico C (2006) The complete mitochondrial genome of the whiting, *Merlangius merlangus*, and the haddock, *Melanogrammus aeglefinus*: a detailed genomic comparison among closely related species of the Gadidae family. *Gene* 383:12–23
- Sarkar M, Das S, Bandyopadhyaya A, Ray K, Chaudhuri K (2005) Upregulation of human mitochondrial NADH dehydrogenase subunit 5 in intestinal epithelial cells is modulated by *Vibrio cholerae* pathogenesis. *FEBS Lett* 579:3449–3460
- Satoh TP, Miya M, Endo H, Nishida M (2006) Round and pointed-head grenadier fishes (Actinopterygii: Gadiformes) represent a single sister group: evidence from complete mitochondrial genome sequences. *Mol Phylogenet Evol* 40:129–138
- Sbisa E, Tullo A, Nardelli M, Tanzariello F, Saccone C (1992) Transcription mapping of the oriL region reveals novel precursors of mature RNA species and antisense RNAs in rat mitochondrial genome. *FEBS Lett* 296:311–316
- Sbisa E, Tanzariello F, Reyes A, Pesole G, Saccone C (1997) Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205:125–140
- Slomovic S, Laufer D, Geiger D, Schuster G (2005) Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. *Mol Cell Biol* 25:6427–6435
- Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (and other methods). Version 4. Sinauer and Associates, Sunderland
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Teletchea F, Laudet V, Hanni C (2006) Phylogeny of the Gadidae (sensu Svetovidov, 1948) based on their morphology and two mitochondrial genes. *Mol Phylogenet Evol* 38:189–199
- Temperley RJ, Richter R, Dennerlein S, Lightowlers RN, Chrzanowska-Lightowlers ZM (2010a) Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* 32:301
- Temperley RJ, Wydro M, Lightowlers RN, Chrzanowska-Lightowlers ZM (2010b) Human mitochondrial mRNAs-like members of all families, similar but different. *Biochim Biophys Acta* 1797:1081–1085
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882
- Tullo A, Tanzariello F, D'Erchia AM, Nardelli M, Papeo PA, Sbisa E, Saccone C (1994) Transcription of rat mitochondrial NADH-dehydrogenase subunits: presence of antisense and precursor RNA species. *FEBS Lett* 354:30–36
- Ursvik A, Breines R, Christiansen JS, Fevolden S-E, Coucheron DH, Johansen SD (2007) A mitogenomic approach to the taxonomy of pollocks: *Theragra chalcogramma* and *T. finnmarchica* represent one single species. *BMC Evol Biol* 7:86
- Valverde JR, Marco R, Garesse R (1994) A conserved heptamer motif for ribosomal RNA transcription termination in animal mitochondria. *Proc Natl Acad Sci USA* 91:5368–5371
- Wallace DC (2007) Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. *Annu Rev Biochem* 76:781–821
- Xiao S, Scott F, Fierke CA, Engelke DR (2002) Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes. *Annu Rev Biochem* 71:165–189
- Yanagimoto T, Kitamura T, Kobayashi T (2004) Complete nucleotide sequence and variation of mitochondrial DNA from 10 individuals of walleye Pollock, *Theragra chalcogramma*. *Fish Sci* 70:885–895
- Zardoya R, Perez-Martos A, Bautista JM, Montoya J (1995) Analysis of the transcription products of the rainbow trout (*Oncorhynchus mykiss*) liver mitochondrial genome: detection of novel mitochondrial transcripts. *Curr Genet* 28:67–70