We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Expanding the Coding Potential of Vertebrate Mitochondrial Genomes: Lesson Learned from the Atlantic Cod

Tor Erik Jørgensen and Steinar Daae Johansen

Abstract

Vertebrate mitochondrial genomes are highly conserved in structure, gene content, and function. Most sequenced mitochondrial genomes represent bony fishes, and that of the Atlantic cod (Gadus morhua) is the best characterized among the fishes. In addition to the well-characterized 37 canonical gene products encoded by vertebrate mitochondrial genomes, new classes of gene products representing peptides and noncoding RNAs have been discovered. The Atlantic cod encodes at least two peptides (MOTS-c and humanin (HN)), two long noncoding RNAs (lncCR-L and lncCR-H), and a number of small RNAs. Here, we review recent research in the Atlantic cod focusing on putative mitochondrial-derived peptides, the mitochondrial transcriptome, and noncoding RNAs.

Keywords: Gadus morhua, long noncoding RNA, mitogenome, mitochondrial-derived peptide, mitochondrial transcriptome, mitochondrial small RNA, mtDNA

1. Introduction

The mitochondrial genome (mitogenome) is highly conserved among vertebrates [1]. All species investigated to date contain mitogenomes encoding the same 37 canonical gene products, organized in a highly similar gene order in most species. Complete mitogenome sequences have been determined from almost 5000 vertebrate species, where about 50% is represented by the bony fishes [2].

The Atlantic cod (Gadus morhua) is a benthopelagic fish in the Gadidae family, belonging to the order of Gadiformes [3, 4]. The 16.7 kb circular mitogenome was one of the first to be
Figure 1. The Atlantic cod mitochondrial genome. (A) Circular map presenting gene content and organization. The mitochondrial genome harbors 13 protein-coding genes (light blue), 2 rRNA genes (yellow), 22 tRNA genes (red), and noncoding regions (gray). CR, control region; \( H_1 \) and \( H_2 \), H-strand promoters; LSP, L-strand promoter; \( O_H \) and \( O_L \), origins of heavy- and light-strand replication, respectively; HTR, heteroplasmic tandem repeat; T–P spacer, intergenic noncoding spacer. tRNA genes are indicated by the standard one-letter symbols for amino acids. All genes are H-strand encoded, except Q, A, N, C, Y, S, E, P, and ND6 (L-strand encoded). mtSSU and mtLSU, mitochondrial small- and large-subunit rRNA genes; ND1–ND6, NADH dehydrogenase subunit 1–6; COI–COIII, cytochrome c oxidase subunit I–III; Cyt b, cytochrome b; ATP6 and ATP8, ATPase subunit 6 and 8. (B) Schematic view of the OxPhos complexes embedded in the inner mitochondrial membrane. ATP is generated by oxidative phosphorylation. The mitochondrial genome encodes 13 of the approximately 85 subunits, belonging to complex I (blue), complex III (orange), complex IV (green), and complex V (yellow).
completely sequenced from a fish species [5–7]. Atlantic cod possesses the same mitogenome organization as most vertebrate species, including that of humans and vertebrate model systems like mouse, rat, *Xenopus*, and zebrafish (Figure 1A).

Among the canonical gene products encoded by the Atlantic cod mitogenome, 13 represent hydrophobic proteins essential for oxidative phosphorylation (OxPhos), two are ribosomal RNAs (rRNAs) of the mitochondrial ribosome, and 22 are transfer RNAs (tRNAs) necessary for mitochondrial translation. The OxPhos system consists of five large protein complexes embedded in the inner mitochondrial membrane. However, only 13 of the approximately 85 OxPhos proteins are encoded by the mitogenome (Figure 1B) [8].

Both strands (H- and L-strands) have coding potential (Figure 1A). Most mitochondrial genes are encoded by the H-strand and include the small and large subunit rRNAs (mtSSU rRNA and mtLSU rRNA), 14 tRNAs, and 12 protein-coding genes. The L-strand, however, encodes only eight tRNAs and one protein. The control region (CR), located between the genes of tRNA^Phe^ and tRNA^Pro^, is the major noncoding region in the mitogenome and constitutes approximately 1000 bp in Atlantic cod [7, 9]. The CR harbors the genetic control elements for H-strand replication origin (OriH), the transcription initiation sites for H- and L-strands, as well as the displacement loop (D-loop) located between OriH and the termination associated sequence (TAS) [7, 9, 10]. Furthermore, a 30-bp spacer located between the genes of tRNA^Asp^ and tRNA^Cys^ contains the origin of L-strand synthesis. OriL appears functionally conserved in most vertebrates [11, 12], including the Atlantic cod [5].

Hallmarks of Atlantic cod mitogenomes are the noncoding intergenic T–P spacer, and the heteroplasmic tandem repeat (HTR) array at the 5’ domain of CR (Figure 1A). The 74-bp Atlantic cod T–P spacer [5, 13], located between the tRNA^Phe^ and tRNA^Pro^ genes, represents an evolutionary preserved feature present in all gadiform species [10, 13]. The T–P spacer is variable in sequence and size among gadiforms but still harbors two conserved 17-bp sequence motifs forming potential hairpin structures at the RNA level [10]. The HTR array consists of a 40-bp sequence motif usually present in two to five copies within an individual [5, 14, 15] and thus results in size heterogeneity and heteroplasmcy of Atlantic cod mitogenomes. Here, we review recent developments in the characterization of Atlantic cod mitogenomes with focus on interindividual sequence variation, mitochondrial transcriptome, noncoding RNAs, and putative mitochondrial-derived peptides.

2. Sequence variation among Atlantic cod mitochondrial genomes

Complete mitogenome sequences have been obtained from approximately 200 specimens representing major ecotypes and geographic locations of Atlantic cod. In one study, based on SOLiD deep sequencing, we performed pooled sequencing of 44 specimens from each of the migratory northeast arctic cod (NA) and the stationary Norwegian coastal cod (NC) [16]. The sequencing represented more than 1100 times mitogenome coverage of each ecotype and 25 times coverage of each individual. We found a total of 365 SNP loci in the dataset, where 121 SNPs were shared between the ecotypes. One hundred fifty-one SNPs and ninety-three SNPs...
were specific to NA and NC cod, respectively. From the dataset we determined the mitochondrial substitution rate to be 14 times higher compared to that of the nuclear genome [16, 17].

More recently we analyzed 156 Atlantic cod mitogenomes at the individual level [18], including 32 specimens previously reported by Carr and Marshall [19]. We found 1034 SNPs in total among the sequences, which were not evenly distributed throughout the mitogenome. The ND2 gene (Complex I) and the COII gene (Complex IV) were the least and most conserved, respectively, among the protein-coding genes. Furthermore, rRNA and tRNA genes showed a significantly lower density of overall SNPs per site compared to protein-coding genes. Thus, the Atlantic cod mitogenome follows a similar pattern of conservation as seen for other vertebrates like zebrafish and human [20–23] and corroborates the observation that mutation rate constrains in vertebrate mitogenomes appear linked to the position of genes in relation to OriH and OriL [24, 25].

The noncoding regions of the Atlantic cod mitogenome showed a mosaic pattern of sequence conservation. Whereas the OriL and the central domain of CR were almost invariant among specimens, the T–P spacer and 5′ domain of CR contain significant sequence variation [7, 10, 13, 18]. The 74-bp T–P spacer was found to contain 16 variable sites and 26 haplotypes among 225 specimens assessed, including a 29-bp sequence duplication in three individuals [10]. Similarly, the 5′ domain of CR was the most variable region within the mitochondrial genome (more than three times that of average substitution rate). The elevated sequence variation was due to hot-spot substitution sites, homopolymeric heterogeneity, and the HTR array [18].

3. Mitochondrial-derived peptides

Vertebrate mitogenomes have the potential of encoding several short peptides (mitochondrial-derived peptides (MDPs)) [26–28]. The best characterized peptides among the MDPs are MOTS-c and humanin (HN). Genes coding for MOTS-c and HN are found as small open reading frames within the mitochondrial small subunit (mtSSU) and large subunit (mtLSU) ribosomal DNA, respectively [29, 30]. Studies in mammals indicate that MDPs are circulating signaling molecules with a number of proposed roles. While HN is involved in cellular stress resistance, apoptosis, and metabolism [29, 31–34], MOTS-c apparently represents an MDP hormone that regulates metabolic homeostasis and insulin sensitivity [30, 35].

The Atlantic cod open reading frames encoding MOTS-c and HN were identified at the exact same locations as in human, within the domain 3′M and domain IV of the mtLSU rRNA and mtSSU rRNA, respectively (Figure 2A and B). Comparative analysis revealed MOTS-c and HN to be invariant among Atlantic cod specimens [18] and well conserved between Atlantic cod and human (Figure 2C). Here, 8 of 16 amino acid residues in MOTS-c and 13 of 21 amino acid residues of HN were shared. Furthermore, when comparing gadiform species representing seven diverse families, we noted 10 of 16 and 15 of 21 amino acid residues to be shared in MOTS-c and HN, respectively (Figure 2C). The conserved features seen between gadiform species and human suggest related MDP functions.
Figure 2. Putative mitochondrial-derived peptides in Atlantic cod. (A) Secondary structure diagram of the Atlantic cod mtSSU rRNA domain 3’M coding for the putative MOTS-c peptide (red letters). (B) Secondary structure diagram of the Atlantic cod mtLSU rRNA domain IV coding for the putative HN peptide (red letters). (C) Alignment of MDP (MOTS-c and HN) sequences from seven gadiform species representing different families (gm, Gadus morhua, Gadidae, HG214359; Li, Lota lota, Lotidae, AP004412; mm, Merluccius merluccius, Merlucciidae, FR751402; Ec, Enchelyopus cimbrius, Phycidae, AJ315624 and FJ215015; tm, Trachyrincus murrani, Macrouridae, AP008990; Bn, Bregmaceros nectabanus, Bregmacerotidae, AP004409; Pj, Physiculus japonicus, Moridae, AP004409; Hs, Homo sapiens, NC_012920). Stars above and below the alignment represent conserved residues among gadiforms and between gadiforms and human, respectively.
4. Mitochondrial transcriptome

Vertebrate mitochondrial transcriptomes have mainly been studied in human cells and tissues [36, 37]. Mature mitochondrial RNAs are generated from three polycistronic transcripts initiated within CR from two H-strand promoters (HSP_1 and HSP_2) and a single L-strand promoter (LSP) (Figure 3A) [36, 38–40]. The HSP_1-specific transcript is highly abundant and generates mtSSU rRNA, mtLSU rRNA, as well as tRNA^{Val} and tRNA^{Phe} [41, 42]. HSP_1-specific tRNAs have recently been proposed to perform a second role as a mitochondrial rRNA, substituting the lacking 5S rRNA in vertebrate mitochondrial ribosomes [43, 44]. While tRNA^{Val} appears associated with the mitochondrial ribosomes in human and rat, tRNA^{Phe} has been identified in porcine and bovine [45].

Ten H-strand-specific mRNAs are posttranscriptionally processed from the HSP_2 transcript, together with 13 tRNAs and the two rRNAs (Figure 3A) [46]. Most HSP_2 mRNAs are monocistronic, but two of the mRNAs are bicistronic (ND4/4 L mRNA and ATPase8/6 mRNA). Finally, the L-strand-specific transcript gives rise to the ND6 mRNA and eight tRNAs (Figure 3A).

4.1. Atlantic cod mitochondrial mRNAs

Similar to that of human cells, 11 mature mRNAs were readily expressed from the Atlantic cod mitogenome [47]. There are, however, some minor differences in mitochondrial mRNA maturation and modification between human and Atlantic cod. Mapping of the 5’ ends in mitochondrial mRNAs by pyrosequencing revealed that 10 of the 11 mRNAs contain no, or very short (1–2 nt), 5’ untranslated regions (UTRs) [47]. The only exception is the 5’ UTR of the COII mRNA, which contained a short hairpin structure. In Atlantic cod and all other Gadidae species, this hairpin structure is capped by a GAAA tetra-loop (Figure 3B) [47]. GAAA tetra-loops are known to frequently participate in long-range RNA:RNA tertiary interactions [48].

Most Atlantic cod mRNAs lack 3’ UTRs, but the COI mRNA has a 3’ UTR of 76 nt corresponding to the complete mirror sequence of tRNA^{Ser(UCA)} (Figure 3B) [47]. A very similar 3’ UTR (72 nt) has been reported in the human COI mRNA [49], indicating a conserved role in vertebrates. The 3’ UTR of the ND5 mRNA is highly variable in length in vertebrates but is lacking completely in Atlantic cod [40, 47]. However, the closely related Gadidae species Pollachius virens (Saithe) contains an ND5 mRNA 3’ UTR of 16 nucleotides [47]. In humans, mitochondrial mRNAs contain short polyA tails of 40–50 adenosines at their 3’ ends [40, 45]. PolyA tails were identified in all mRNAs, except for ND6 mRNA [40], and seven UAA termination codons were created in the human mitochondria by polyA posttranscriptional editing [50]. Similarly, all mitochondrial mRNAs (except the ND6 mRNA) were found to be polyadenylated in Atlantic cod, and six UAA termination codons were generated by polyA addition [47].

4.2. Atlantic cod mitochondrial structural RNAs

The 22 mitochondrial tRNAs were found to be highly conserved in Atlantic cod, both in structure and sequence [5, 18], and some tRNAs (tRNA^{Phe}, tRNA^{Ser(UCA)}, tRNA^{Ser(AGY)}, and tRNA^{Cys})
Figure 3. The mitochondrial transcriptome in Atlantic cod. (A) Schematic map of mitochondrial ribosomal RNA, messenger RNA, and long noncoding RNA generated from HSP1, HSP2, and LSP transcripts. mtSSU rRNA and mtLSU rRNA, mitochondrial small- and large-subunit ribosomal RNA (yellow boxes); ND1, ND2, ND3, ND4L/ND4, ND5, and ND6, NADH dehydrogenase subunit mRNAs; COI, COII, and COIII, cytochrome c oxidase subunit mRNAs; A8/A6, ATPase subunit bicistronic mRNA; Cyt B, cytochrome b mRNA (all mRNAs indicated by blue boxes); IncCR-H and IncCR-L, long noncoding RNAs (orange boxes). (B) 3′ untranslated region (UTR) and 5′ UTR in COI and COII mRNAs, respectively. Translation initiation codons (GUG and AUG) and termination codons (UAA) are indicated by green and red circles, respectively. The 3′ UTR of COI mRNA contains a mirror tRNA motif, and the 5′ UTR of COII mRNA contains a GAAA tetra-loop hairpin motif.

were invariant in the 200 specimens investigated. SOLiD deep sequencing confirmed a non-template CCA addition at the 3′ ends of tRNAs (our unpublished results). Thus, mitochondrial tRNA processing and probably modification are highly similar in human and Atlantic cod [47].
The annotated mtSSU rRNA and mtLSU rRNA genes in Atlantic cod are 952 and 1664 bp, respectively [7]. The corresponding rRNAs are highly conserved within the species [18] and well conserved between different fish species [7, 51]. The 5′ and 3′ ends of Atlantic cod mitochondrial rRNAs have been precisely mapped using different approaches. Primer extension and pyrosequencing confirmed the 5′ ends to correspond to the annotated features based on comparative sequence alignments [47, 51]. The 3′ ends were mapped by pyrosequencing and by RNA ligation sequencing [51]. Interestingly, non-template adenosines were added at both rRNAs. Whereas the 3′ end of mtSSU rRNA was found to be homogenous and mono-adenylated, the corresponding end of mtLSU rRNA was heterogeneous and oligo-adenylated [51]. The observed mtLSU rRNA heterogeneity is consistent with the notion that mitochondrial rRNAs are transcribed and processed from two different precursor RNAs, the HSP$_1$ and HSP$_2$ primary transcripts (Figure 3A).

5. Mitochondrial noncoding RNAs

In addition to the canonical mitochondrial genes and the newly proposed MDPs, vertebrate mitogenomes encode several noncoding RNAs [36]. The first discovered mitochondrial long noncoding RNA (lncRNA) was the human L-strand-specific 7S RNA (lncCR-L) [52, 53]. At least eight vertebrate mitochondrial lncRNAs have now been proposed and characterized [54]. Two lncRNAs correspond to the H-strand and L-strand of CR (lncCR-H and lncCR-L) [10, 18, 47, 52, 55, 56], one is an antisense chimera to partial regions of the CytB and COI mRNAs (LIPCAR) [57–59], three are mRNA antisense RNAs (IncND5, IncND6, and IncCytB) [60], and two are chimeric RNAs that involve sense and antisense mtLSU rRNAs (SncmtRNA and ASncmtRNA) [61–63]. So far, LIPCAR, rRNA chimeras, and IncCR-H have been associated with human diseases [56, 57, 61, 63–66]. There are apparently a large number of small non-coding RNAs (mitosRNAs) generated from vertebrate mitochondrial transcripts [36, 67–69]. None of these mitosRNAs have been assigned to a specific function fund on experimental evidence. However, in a recent study by Riggs and Podrabsky [70], mitosRNAs were associated to a hypoxia stress response in killifish embryos.

5.1. Atlantic cod mitochondrial long noncoding RNAs

Two lncRNAs (lncCR-H and lncCR-L) have been identified and investigated in Atlantic cod mitochondria (Figure 4) [10, 18, 47]. Both lncRNAs were found to be polyadenylated but transcribed from opposite strands within the CR [10]. We showed that the Atlantic cod lncCR-L has a mutation rate and an expression level corresponding to that of Complex I mRNAs [10, 18, 47]. The lncCR-L apparently corresponds to the 7S RNA in human mitochondria [52], and recently we showed that lncCR-L is differentially expressed in a human cancer-matched cell line pair [56].

The lncCR-H was found to be highly variable in sequence and structure, both between and within Atlantic cod specimens [10, 18]. A schematic overview of the lncCR-H RNA is presented in Figure 4. Here, the noncoding T–P spacer is present at the 5′ end and includes two potential RNA hairpin structures. The T–P spacer domain is followed by a mirror tRNA$^{Pro}$, before entering the HTR array motifs. The HTR copy numbers vary between 2 (80 bp)
and more than 8 (>320 bp) [5, 14, 15, 18], rendering IncCR-H highly variable in size. Finally, IncCR-H terminates in a short polyA tail at TAS. Thus, IncCR-H has apparently no fixed length in Atlantic cod mitochondria and varies in size between approximately 300 and 500 nt. Interestingly, the TAS motif consists of a perfect palindromic sequence motif (TTTATACAT ATGTATAAA). We found IncCR-L to terminate with a polyA tail at the same site as IncCR-H but on the opposite strand [10].

5.2. Atlantic cod mitochondrial small RNAs

The Atlantic cod mitogenomes express a number of small RNAs, revealed by SOLiD small RNA sequencing experiments (our unpublished results). Here, the majority of mitosRNA was identified as mitochondrial tRNA-derived fragments (tRFs; see [69, 70]). Interestingly, most Atlantic cod mitochondrial tRFs correspond to H-strand tRNAs, and some tRFs were differentially expressed during early developmental stages (our unpublished results). Many of the same tRF species detected in Atlantic cod have recently been noted in rainbow trout egg cells [69] and in killifish embryos [70], suggesting a conserved feature at least among some bony fishes.

The SOLiD experiments also detected several abundant small RNAs mapping to the mitochondrial CR [17]. We found three small RNA candidates generated from lncCR-L, suggesting this lncRNA to be a precursor for mitosRNA (Figure 4). Similarly, two mitosRNA were generated from IncCR-H, one corresponded to a pyrimidine-rich motif and the other to tRF-1 derived from tRNAThr (Figure 4). What functions these small RNAs may serve in the

![Figure 4. Schematic view of CR and corresponding noncoding RNAs in Atlantic cod. tRNA genes (tRNAThr, tRNAPro, tRNAPhe), T–P spacer, HTR (heteroplasmonic tandem repeat array), TAS (termination associated sequence), and CSB2 (conserved sequence box 2) are indicated. The H-strand-specific IncCR-H is located at the 5′ domain of CR and is the precursor of two enriched small RNAs (above CR map). The L-strand-specific IncCR-L is located at the central domain of CR and is the precursor of three enriched small RNAs (below CR map).](http://dx.doi.org/10.5772/intechopen.75883)
mitochondria are not currently known, but we speculate that regulatory roles related to transcription elongation, mtDNA replication, or ribosome functions are likely.

6. Concluding remarks

The mitochondrial gene content and organization are highly conserved between Atlantic cod and human and strongly support a common functional platform. Similarly, the mitochondrial transcripts generating canonical mRNAs and structural RNAs are surprisingly similar. What about the newly proposed MDPs and noncoding RNAs? Are there any lineage-specific differences? Research is still in its infancy, but recent findings suggest conservations between fish and mammals. More experimental studies in Atlantic cod and model systems like zebrafish are highly encouraged, including investigations of the fascinating mitochondrial swinger RNAs [24, 71, 72]. Mitochondrial-derived noncoding RNAs need to be profiled and further investigated in adult tissue types during normal and stress conditions, as well as at various developmental stages. A first step could be to study the intracellular location by in situ RNA hybridization and then ask if the noncoding RNAs are confined to the mitochondrial compartment or exported to the cytoplasm or other cellular compartments. Our studies in Atlantic cod indicate that at least two of the mitochondrial IncRNAs may serve as precursors for small RNAs. We conclude that vertebrate mitogenomes encode a significant number of gene products in addition to the 37 canonical OxPhos proteins, rRNAs, and tRNAs.

Acknowledgements

We thank the current and former members of the mitochondrial DNA research team at the Genomics Group, Nord University, for the discussion and support. This work was supported by Nord University, Bodø, Norway.

Conflict of interest

The authors declare that they have no conflict of interest.

Author details

Tor Erik Jørgensen and Steinar Daae Johansen*

*Address all correspondence to: steinar.d.johansen@nord.no

Faculty of Biosciences and Aquaculture, Genomics Group, Nord University, Bodø, Norway
References


[18] Jørgensen TE. Molecular and evolutionary characterization of the Atlantic cod mitochondrial genome [PhD thesis]. Norway: Nord University; 2018


[33] Paharkova V, Alvarez G, Nakamura H, Cohen P, Lee K-W. Rat Humanin is encoded and translated in mitochondria and is localized to the mitochondrial compartment where it regulates ROS production. Molecular and Cellular Endocrinology. 2015;413:96-100. DOI: 10.1016/j.mce.2015.06.015

[34] Capt C, Passamonti M, Breton S. The human mitochondrial genome may code for more than 13 proteins. Mitochondrial DNA Part A. 2016;27:3098-3101. DOI: 10.3109/19401736.2014.1003924


[40] Temperley RJ, Wydro M, Lightowlers RN, Chrzanowska-Lightowlers ZM. Human mitochondrial mRNAs-like members of all families, similar but different. Biochimica et Biophysica Acta. 1797;2010:1081-1085. DOI: 10.1016/j.bbabio.2010.02.036


[59] Zhang Z, Gao W, Long QQ, Zhang J, Li YF, Liu DC, Yan JJ, Yang ZJ, Wang LS. Increased plasma levels of lncRNA H19 and LIPCAR are associated with increased risk of coronary artery disease in a Chinese population. Scientific Reports. 2017;7:7491. DOI: 10.1038/s41598-017-07611-z


[71] Seligmann H. Sharp switches between regular and swinger mitochondrial replication: 16S rDNA systematically exchanging nucleotides $A \leftrightarrow T + C \leftrightarrow G$ in the mitogenome of *Kamimuria wangi*. Mitochondrial DNA Part A. 2016;27:2440-2446. DOI: 10.3109/19401736.2015.1033691
