

# 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

Severine Roques <sup>a</sup>, Clive J. Fox <sup>b</sup>, Maria I. Villasana <sup>a</sup>, Ciro Rico <sup>a, </sup>

<sup>a</sup> Estación Biológica de Doñana, CSIC, Pabellón del Perú, 41013 Sevilla, Spain

<sup>b</sup> The Centre for Environment, Fisheries and Aquaculture Science, Pakefield Road, Lowestoft, Suffolk, NR33 OHT, England, United Kingdom

## Abstract

We determined the first complete mitochondrial DNA (mtDNA) sequences for the whiting (*Merlangius merlangus*, family Gadidae, order Gadiformes) and the haddock (*Melanogrammus aeglefinus*, family Gadidae, order Gadiformes). The entire mitogenomes were amplified and sequenced by primer walking using newly designed specific internal primers. Lengths were 16,569 and 16,585 bases for whiting and haddock respectively, lengths which lie within the range of previously reported gadiform sequences from Atlantic cod (*Gadus morhua*, 16,696 bases) and walleye pollock (*Theragra chalcogramma*, 16,570 bases). Gene arrangement in both species conformed to the order seen in most vertebrate mitochondrial genomes. We identified a long intergenic spacer located between the tRNA<sup>Thr</sup> and tRNA<sup>Pro</sup> genes (of 100 and 70 bp long for whiting and haddock, respectively), as previously described for other species of the order Gadiformes. Using nucleotide and amino acid divergence data of four complete gadoid mitogenomes (*M. merlangius*, *M. aeglefinus*, *G. morhua* and *T. chalcogramma*), we examined in detail the relative mtDNA mutation patterns across genes and among Gadidae species and tested for the performance of each protein-coding, transfer RNA and ribosomal RNA gene in depicting the expected phylogeny among the four species, as compared with the whole genome dataset. This comparison may be particularly useful in phylogenetic analyses of such a diverse fish family, as well as for the understanding of the patterns of nucleotide substitution of the mtDNA at low levels of divergence.

Keywords: Mitochondrial DNA; Complete sequence; Fish; Protein-coding genes

## 1. Introduction

From the numerous studies, it appears that the optimal resolution of a particular mitochondrial DNA (mtDNA) gene for population or phylogenetic inferences relies on an optimal

combination of the mutation rate of the portion analysed and on the level of polymorphism studied. Because different regions of the mt genome evolve at different rates, certain regions of the mtDNA have been commonly targeted for particular types of study, e.g. phylogeny, phylogeography or population structure (Avice, 1986). Thus, in phylogenetic studies, one usually assumes that a single portion of mtDNA (generally 300–600 base pairs) is representative of both the complete molecule and the maternal evolutionary relationships of the organism studied. However, increasing evidence from gene-comparison studies has demonstrated that the phylogenetic performance may vary among different genes. A single gene may not be representative of the whole mitogenome and can provide a misleading inference of the true inter-specific evolutionary relationships (Cao

---

Abbreviations: A, adenosine; aa, amino acid (s); C, cytosine; CR, control region; CSB, conserved block sequence; DHU, dihydrouracil; G, guanosine; mt, mitochondrial; MRP, mean ranking position; NC, non-coding; ORF, open reading frames; PCG, protein-coding gene (s); rRNA, ribosomal RNA; T, thymidine; TAS, termination-associated sequence; tRNA, transfer RNA; U, uridine.

 Corresponding author. Tel.: +34 954232340; fax: +34 954621125.  
E-mail address: [ciro@ebd.csic.es](mailto:ciro@ebd.csic.es) (C. Rico).



et al., 1994; Meyer, 1994; Cummings et al., 1995; Russo et al., 1996; Zardoya and Meyer, 1996; Miya and Nishida, 2000). Moreover, the relationship between gene mutation rate and our ability to recover the true evolutionary relationships may not be straightforward. Obtaining complete genomic sequences from extensive taxonomic sampling may overcome these problems, thus improving estimates of tree topology and dates of divergence among clades (Pollock et al., 2000). In addition, these data may subsequently aid in choosing the appropriate genetic marker(s) for a given level of divergence (e.g. shallow or deep phylogenies or population genetics) among related species.

The large amount of recently available sequence data indicates that fish mtDNA is structurally very similar to that of other vertebrates with 22 tRNA genes, 13 protein-coding genes and a main non-coding region (control region) with common features such as TAS (termination-associated sequences) and CSB (conserved block sequences) (Inoue et al., 2000; Delarbre et al., 2001; Lee et al., 2001; Miya et al., 2001; Doiron et al., 2002; Miya et al., 2003; Kim et al., 2004; Machado et al., 2004). Among approximately 140 complete sequences, however, nine patterns of gene rearrangement were found in euteleosts (Miya and Nishida, 1999; Inoue et al., 2001; Miya et al., 2001, 2003). Also, Miya and Nishida (1999) reported the first example of tRNA gene rearrangement in bony fishes between the ND6 and the control region (ND6–cytb–tRNA<sup>Glu</sup>–tRNA<sup>Pro</sup>–tRNA<sup>Thr</sup>–control region), changing the “conserved” status of fish mt genome. Although complete genomes have been particularly helpful in resolving phylogenetic relationships, subsequent sequencing and analyses may still be necessary for resolving persistent controversies over higher-level relationships within teleosts which are the most diversified group of all vertebrates.

The order Gadiformes comprises more than 450 species in 21 genera, arranged in 11 families: Bregmacerotidae, Euclichthyidae, Gadidae, Lotidae, Macrouridae, Macruronidae, Melanonidae, Merlucciidae, Moridae, Muraenolepidae and Muraenolepididae (Cohen et al., 1990). The family Gadidae is a large family with 15 known species in 12 genera and includes a large number of commercially important species, such as saithe, pollock, cod, haddock and whiting. Complete mtDNA sequences have been reported for six species of Gadiformes [Lota lota, family Lotidae, Caelorinchus kishineryei, family Macrouridae, Pysiculus japonicus, family Moridae Bregmaceros nectabanus, family Bregmacerotidae, Gadus morhua (Johansen and Bakke, 1996) and Theragra chalcogramma, family Gadidae (Yanagimoto et al., 2004; Miya et al., 2003)]. Further, special features of Gadiformes mt genomes have also been reported. For example, Miya et al. (2001) reported the presence of a unique mitogenome *B. nectabanus*, characterized by many rearrangements, such as the unusual position of the cytb and some tRNAs (tRNA<sup>Ala</sup>, tRNA<sup>Asn</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Ser</sup>, tRNA<sup>Glu</sup>, tRNA<sup>Thr</sup>) genes. Also, Bakke et al. (1999) determined the unique presence of intergenic spacers between the tRNA<sup>Thr</sup> and tRNA<sup>Pro</sup> in the mt DNA of *G. morhua*, *T. chalcogramma*, *Boreogadus saida*, *Melanogrammus aeglefinus*, *Micromesistius poutassou*, *Brosme brosme*, *Enchelyopus cimbricus* and *Gaidropsarus argentatus*.

Among the gadiform taxa, genetic variation of the cytochrome b gene has proved very useful for inter-specific evolutionary relationships, but has given incoherent results within some genera of the family Gadidae for which the monophyly is still debated (Carr et al., 1999; Moller et al., 2002). Lack of genetic resolution may be the result of the low performance of the gene analysed to describe several speciation events within a relatively short period of time (Carr et al., 1999). However, the recent alignment of ten gadiform mt ribosomal and cytochrome b genes has clarified some of the unresolved phylogenetic relationships, corroborating that the Gadidae is the most derived group, with the most recent genus being *Gadus*, *Boreogadus*, *Merlangius*, *Melanogrammus* and *Pollachius* (Bakke and Johansen, 2005). Although at the intra-specific level mtDNA genetic variation has been extensively studied in Atlantic cod (*G. morhua*) (Carr and Marshall, 1991; Carr et al., 1995; Arnason and Palsson, 1996; Arnason et al., 1998, 2000; Sigurgislason and Arnason, 2003; Arnason, 2004) very little is known on most other commercial gadoid species.

The whiting (*Merlangius merlangus*) and the haddock (*M. aeglefinus*) belong to the family Gadidae and are among the most abundant and commercially important gadoid fish species of the North Atlantic. There are marked differences in geographic distribution with haddock widely distributed across both sides of the Atlantic. In contrast, whiting occur only on the eastern shelves but penetrate further south reaching the Mediterranean Sea (Wheeler, 1978). Based on mitochondrial cytochrome b and cytochrome oxidase I genes, previous phylogenetic studies concluded that both species are closely related and sister taxa to the more recently diverged group of cod-like and pollock species (Carr et al., 1999; Moller et al., 2002; Bakke and Johansen, 2005). At the intra-specific level, little resolution was found in the haddock mt DNA variation (12 restriction endonucleases), either because the fragment was too short to contain sufficient variability or because the gene was not necessarily suitable for the purpose of this study (Zwanenburg et al., 1992).

Our study describes for the first time, the complete mtDNA sequences of whiting and haddock. Gene arrangement and basics statistics are given providing information on the extent of polymorphism and divergence between these closely related species. The mtDNA genomes of these species will be an important source of sequence information for improving primer selection and design and for the identification of diagnostic regions suitable for species-specific identification, particularly of early stage eggs from plankton surveys (Taylor et al., 2002; Fox et al., 2005). Using nucleotide and amino acid divergence data of the four complete gadoid mitogenomes so far sequenced (whiting, haddock, this study, and Atlantic cod and walleye pollock, from GenBank), we examined the relative mtDNA mutation patterns across genes and among gadoid species. In addition, we tested for the performance of each protein-coding gene (PCG), the combined transfer RNA genes and the two ribosomal RNA subunits in depicting the expected phylogeny among the four species, as compared to the whole genome and the combined PCG dataset. This should be of particular use for phylogenetic studies in such a diverse species family, as well as

for the understanding of the patterns of nucleotide substitutions of the mtDNA, at low levels of divergence.

## 2. Materials and methods

### 2.1. Sample and DNA extraction

A small portion of gill or muscle was removed from freshly caught adult whiting and haddock from the North Sea and stored in 99% ethanol at room temperature. Fish were captured during fisheries research cruises and the fish were identified by experienced fisheries scientists. Prior to DNA extraction, tissues were washed in distilled water and dried on tissue paper to remove excess ethanol. Total genomic DNA was extracted by standard proteinase K digestion, high-salt chloroform extraction and DNA precipitation, as outlined in Rico et al. (1992).

### 2.2. PCR and sequencing by primer walking

The same sequencing strategy was used for both species. The total length (~ 16.0 kb) of mtDNA was amplified by polymerase chain reaction (PCR). Initially, two regions of the mt genome were obtained, in the cytb (primers GluDG and CB3H) and ATPase (primers L8 331 and CO3.2) genes using primers that were designed from the identity observed among fishes mtDNA (E. Bermingham, unpublished data) (see Appendix 1). PCR amplification was performed in 50 µl reaction containing 2 mM MgCl<sub>2</sub>, each dNTP at 2 µM, each primer at 10 pM, 0.5 U of BioTaq (Bioline) and the supplied buffer at 1× concentration. Amplifications were performed in a Perkin Elmer 9700 thermocycler under the following conditions: for the ~ 900 bp of the ATPase, denaturation at 94 °C for 2 min, 30 cycles of 45 s at 94 °C, 45 s at 58 °C and 1.30 min at 72 °C; for the ~ 850 bp of the cytb, a denaturation step at 94 °C for 5 min, 35 cycles of 45 s at 94 °C, 45 s at 45 °C and 45 s at 72 °C followed by a final step at 72 °C for 2 min. PCR products were purified with Centricon columns and sequenced in both directions for both genes using the BigDye™ terminator sequencing ready kits with AmpliTaq DNA as out lined in Taylor et al. (2002). Primers used were the same as those for the PCRs. Sequencing reactions were resolved on an ABI 377 (Applied Biosystems). Subsequent sequencing was performed using new primers designed with reference to previously-determined sequences (primer walking), in various combinations to amplify contiguous, overlapping segments of the entire mt genome for both species (see primer locations and sequences of these primers in Appendix 1). Complete mt genome sequences of *M. merlangus* and *M. aeglefinus* are deposited in GenBank under accession numbers DQ020496 and DQ020497, respectively.

### 2.3. Sequence analysis

Sequences were aligned with the Seqman (DNASTar Inc., Madison, WI, USA) and Sequencer 4.1.2 (Gene Codes Corp.) programs and improved manually. DNA sequences were compared against the previously sequenced mt genomes of cod, *G. morhua* (Johansen and Bakke, 1996, NC002081) and walleye

pollock, *T. chalcogramma* (Yanagimoto et al., 2004, AB094061) in order to determine the protein-coding and rRNA genes by similarity. Transfer RNA genes (tRNAs) and their respective cloverleaf structures were identified using the program tRNAscan-SE (version 1.1, <http://www.genetics.wustl.edu/eddy/tRNAscan-SE/>, Lowe and Eddy, 1997) using the vertebrate mt codon sequence and setting the cove cut-off score to 1 or manually when tRNAs were not recognized by the program. The base compositional frequencies and the uncorrected p distances for both nucleotide substitutions (transitions and transversions) and amino acid were estimated using MEGA version 2.1 (Kumar et al., 2001). Mitochondrial genomes were analysed for the whole sequences, the combined tRNAs, the ribosomal RNAs (12S and 16S), the control region (CR) and the combined protein-coding genes (PCG). Tamura and Nei evolutionary distances were calculated among the four species for each mtDNA genome portion, using MEGA.

### 2.4. Phylogenetic analyses

The phylogenetic relationships among the four species (*M. merlangus*, *M. aeglefinus*, *G. morhua*, *T. chalcogramma*), were inferred using maximum-parsimony (MP) (MEGA) using heuristic searches (max–mini branch and bound option), neighbor-joining (NJ) (MEGA; based on Tamura and Nei, distance matrices; Tamura and Nei, 1993), and maximum-likelihood (ML) (PAUP\* 4.0b5, Swofford, 1999) using the model parameters of Modeltest 3.06 (Posada and Crandall, 1998) for the combined PCG analyses, and using *L. lota* (accession number, NC004379, Miya et al., 2003), a gadiform of the Lotidae family as outgroup species.

In addition, to evaluate the power of the individual genes in reflecting the extent of divergence among the six species pairs, MP, NJ and ML analyses were conducted for each protein-coding gene, the 12S and 16S ribosomal RNA genes, the combined transfer RNA genes and the whole genome. Analyses with all phylogenetic methods were also performed excluding third codon positions in each gene. Statistical reliability was evaluated using bootstrap (BP) resampling analyses (1000 bootstrap replications) using MEGA. As additional measure, the partitioned Bremer support (PBS) (Baker and DeSalle, 1997) was calculated to evaluate the contribution of each data partition (each protein-coding gene) to the maximum-parsimony tree, using the program Treerot version 2b (Sorenson, 1999). The PBS divides the Bremer support at a given node among various partitions combined in the analysis. Positive values indicate that a given partition provides support for that particular node over the alternative relationships specified in the most parsimonious tree, while a negative value provided contradictory evidence for the relationship found in the tree. A given gene will be considered to perform well, if both the topology and the statistical supports (BP values and PBS) are congruent with the phylogenetic tree of the combined PCG and the whole genomes. It is important to clarify that the magnitude of either the BP or PBS of a particular partition are less relevant than the congruence between them. Generally, high bootstrap values are associated with positive and high PBS values.

Finally, we tested the relative influence of positive selection and neutral evolution on a particular PCG by determining the frequency of aa changes (nonsynonymous, Ka) versus silent base changes (synonymous, Ks). An increase in the ratio of Ka/Ks would then reflect an increased aa substitution rate in that protein, thus relaxed selective constraints. The Ka/Ks ratios for all pairwise combinations of the four Gadidae mtDNA sequences were calculated by using the method of Nei and Gojobori in the DNAsp software (Rozas et al., 2003).

### 3. Results and discussion

#### 3.1. Organization and general structural features of the whiting and haddock mitochondrial genomes

The mtDNA genome of both species was very similar, both in size and content, with the 37 mt structural genes found in all

other vertebrates analysed so far (see Table 1). The overall base composition was similar to that found for the other two gadoid species sequenced so far (cod, *G. morhua*, Johansen and Bakke, 1996 and the walleye pollock, *T. chalcogramma*, Yanagimoto et al., 2004) (percentages for whiting and haddock respectively, Table 2). Base sharing and overlap among tRNA, rRNA and protein-coding genes are also shown in Table 1.

Both genomes contain the 22 tRNA genes typical for metazoans, which are interspersed along the genome (Table 1). The majority of the tRNAs' cloverleaf structures are similar in branch size length. However, 7 tRNAs showed size variability and atypical structures in their T $\psi$ C, DHU and anticodon arms, which were also present in the *G. morhua* and *T. chalcogramma* mitogenomes (Johansen and Bakke, 1996; Yanagimoto et al., 2004) (Fig. 1): tRNA<sup>Val</sup> and tRNA<sup>Cys</sup> have a 2 bp only in the DHU stem, while tRNA<sup>Ile</sup>, tRNA<sup>Lys</sup>, tRNA<sup>Ser</sup>, tRNA<sup>Tyr</sup> showed 3 bp. DHU stem of 3 bp were also recorded in another fish

Table 1

Location of genomic features, size (in bp), start and stop codons of the 13 protein-coding genes (in bold), intergenic nucleotides (ITG, in bp) in the mitochondrial DNAs of *Merlangius merlangius* (Mm: 16,571 bases) and *Melanogrammus aeglefinus* (Ma: 16,587 bases)

Gene	Position number				Size Mm/Ma (bp)	Codon		ITG Mm/ Ma	Strand Mm/ Ma
	Mm		Ma			Mm/Ma Start	Mm/Ma Stop		
	From	To	From	To					
tRNA <sup>Phe</sup>	1	68	1	68	68			0	H
12SrRNA	69	1018	69	1020	950/952			0	H
tRNA <sup>Val</sup>	1019	1090	1021	1092	72			0	H
16SrRNA	1091	2757	1093	2761	1667/1669			0	H
tRNA <sup>Leu</sup>	2758	2831	2762	2835	74			0	H
ND1	2832	3806	2836	3810	975	ATG	TAA	1	H
tRNA <sup>Ile</sup>	3808	3877	3812	3881	70			-1	H
tRNA <sup>Gln</sup>	3877	3947	3881	3951	71			-1	L
tRNA <sup>Met</sup>	3947	4016	3951	4020	70			0	H
ND2	4017	5063	4021	5067	1047	ATG	TAG	-2	H
tRNA <sup>Trp</sup>	5062	5131	5066	5136	70/71			2/1	H
tRNA <sup>Ala</sup>	5134	5202	5138	5206	69			2	L
tRNA <sup>Asn</sup>	5205	5278	5209	5282	74			31	L
tRNA <sup>Cys</sup>	5310	5374	5314	5379	65/66			0	L
tRNA <sup>Tyr</sup>	5375	5441	5380	5446	67			1	L
COI	5443	6996	5448	7006	1554/1559	GTG	TAA	2	H
tRNA <sup>Ser</sup>	6999	7070	7009	7080	72			3	L
tRNA <sup>Asp</sup>	7074	7139	7084	7147	66/64			14	H
COII	7154	7852	7162	7860	699/699	ATG	AGA	-8	H
tRNA <sup>Lys</sup>	7845	7918	7853	7926	74			1	H
ATP8	7920	8087	7928	8095	168	ATG	TAA	-10	H
ATP6	8078	8761	8086	8769	684	ATG	TAA	-1	H
COIII	8761	9546	8769	9554	786	ATG	TAA	0	H
tRNA <sup>Gly</sup>	9547	9614	9555	9622	68			1	H
ND3	9616	9966	9624	9974	351	ATG	TAG	-1	H
tRNA <sup>Arg</sup>	9965	10032	9972	10,041	68/70			0	H
ND4L	10,033	10,332	10,042	10,338	300/297	ATG	TAA	-7	H
ND4	10,326	11,711	10,332	11,714	1386/1383	ATG	AGA	-5	H
tRNA <sup>His</sup>	11,707	11,775	11,710	11,776	69/67			0	H
tRNA <sup>Ser</sup>	11,776	11,843	11,777	11,844	68			2	H
tRNA <sup>Leu</sup>	11,846	11,918	11,847	11,919	73			0	H
ND5	11,919	13,754	11,920	13,758	1836/1839	ATG	TAA	-4	H
ND6	13,751	14,271	13,755	14,275	521	ATG	TAG	0	L
tRNA <sup>Glu</sup>	14,272	14,341	14,276	14,344	70/69			3/4	L
Cytb	14,345	15,505	14,349	15,509	1161	ATG	TAG	-20	H
tRNA <sup>Thr</sup>	15,485	15,555	15,490	15,560	71			70/100	H
tRNA <sup>Pro</sup>	15,625	15,695	15,660	15,730	71			0	L
Control region	15,696	16,569	15,731	16,585	874/857			0	H

Table 2  
Base composition of the whiting (*Merlangius merlangius*) and the haddock (*Melanogrammus aeglefinus*) H strand mitochondrial genomes

	A	C	G	T
Protein	27.1 (26.7)	25.2 (24.8)	14.7 (14.9)	33.0 (33.6)
1st	26.5 (26.6)	25.3 (24.7)	25.2 (25.0)	23.0 (23.6)
2nd	18.0 (18.1)	28.1 (28.0)	13.6 (13.5)	40.4 (40.3)
3rd	36.9 (35.3)	22.3 (21.6)	5.2 (6.2)	35.6 (36.9)
tRNAs	30.2 (30.2)	25.2 (25.4)	20.3 (20.1)	24.3 (24.3)
rRNAs	32.7 (32.8)	23.1 (23.3)	21.0 (20.9)	23.3 (23.0)
Overall	28.9 (28.5)	24.8 (24.6)	16.2 (16.3)	30.1 (30.5)

Whiting (haddock).

species, *R. marmoratus* (Lee et al., 2001). Also, the tRNA<sup>Gly</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Tyr</sup> and tRNA<sup>Val</sup> have 4 bp in the anticodon stem.

There were 13 non-coding (NC) regions, ranging in size from 1 to 874 bp in whiting and from 1 to 857 bp in haddock (see Table 1). The two largest NC regions are found between the tRNA<sup>Thr</sup> and tRNA<sup>Pro</sup> and between tRNA<sup>Pro</sup> and tRNA<sup>Phe</sup>. For both species the major NC region between tRNA<sup>Pro</sup> and tRNA<sup>Phe</sup> corresponds to the control region (CR), that contains apparent homologues of the conserved sequence blocks, CSB-1 (5'-ATAAAGGGTCTTCAAGAGCATAA-3'), CSB-2 (5'-TAAACCCCCCCCCCCCC-3'), CSB-3 (5'-TGCAAA CCCCCCGGAACA-3') in the right domain, and termination-associated sequences (5'-TACATATGTATAAT-3' and 5'-ATCATGAAAAAT-3') in the left domain, as identified in all other fish mtDNA control regions (Lee et al., 1995). The other long NC region between the tRNA<sup>Thr</sup> and tRNA<sup>Pro</sup> genes corresponded to an intergenic spacer previously described for haddock and seven other species of the order Gadiformes. This sequence differs in both length and sequence for whiting, although it contains the two copies of a conserved 17-bp motif, previously identified by Bakke et al. (1999). The presence of this spacer in all species of the family Gadidae sequenced so far is indicative of its relative stability over time and thus, suggests an unidentified functional role in the mtDNA genome of Gadiformes. Gene rearrangements that result in these NC regions have been proposed in birds and salamanders (Quinn, 1997; McKnight and Shaffer, 1997; Zardoya et al., 2003).

### 3.2. Divergence within the Gadidae family

The availability of four complete Gadidae mt DNA genomes allowed us to examine the genetic polymorphism among related species within this family. Total nucleotide divergence (based on Tamura and Nei distances) based on the whole genome ranged from 4.3% between *G. morhua* and *T. chalcogramma* to 9.4% between *T. chalcogramma* and *M. aeglefinus* (Table 3). As previously observed in other comparative genetic studies, the numbers of substitutions (TS and TV) varied considerably among the different mtDNA portions. The concatenated protein-coding genes (PCG) and the control region (CR) are thus the sequences that show the highest divergence values ranging from 5.9% to 13.8% (over 11,472 bp) and from 7% to 18.4% (over 890 bp) respectively, depending on the species pair comparison, while the tRNAs and rRNAs are found to be much more less divergent (Table 3). This is an expected result as

purifying selection is known to constrain nucleotide divergence in both tRNAs and rRNAs due to the secondary structure of these molecules and the specificity of their aa binding sites during translation (Graur and Li, 1999).

All mitochondrial portions, however, highly supported the closest relationship and recent divergence of *G. morhua* and *T. chalcogramma* (Table 3). Also, the phylogenetic relationships among the four species based on the whole mtDNA genomes (for MP, NJ and ML methods) confirmed previous established phylogenies based on smaller mtDNA sections (cytb and COI), with *Merlangius* and *Melanogrammus* genera being closer together than with the more recent group of cod-like and pollock species (including the *Gadus*, *Arctogadus*, *Boreogadus* and *Theragra* genera) (Carr et al., 1999; Moller et al., 2002; Bakke and Johansen, 2002, 2005) (Fig. 4). Based on the whole cytb gene and on the divergence rate of 3.5%/my for synonymous sites (Bermingham et al., 1997), divergence times were estimated between 8.8 and 11 my between *Merlangius* and *Melanogrammus*, 4.6 and 6.1 my between *G. morhua* and *T. chalcogramma* and from 10.4 to 12.4 my among both clades. These results are comparable than those found by Carr et al. (1999) and Bakke and Johansen (2005) based on a portion of the cytb gene and further substantiate the results of a study that examined fossil otoliths, to date the *Melanogrammus*, *Microgadus*, *Theragra* and *Gadus* genera from the Miocene/Pliocene boundary (13 Mya) (Fig. 3 in Nolf and Steurbaut, 1989).

### 3.3. Comparative analysis of the mtDNA protein-coding and RNA genes

PCGs in Fig. 2 are described in the order they are located in the genome. The highest number of substitutions between the most divergent species (*M. aeglefinus* and *G. morhua*) were found in four genes of the NADH enzymatic complex (ND3, ND5, ND4, ND1), and the cytochrome b gene. The highest amino acid changes were found in the ND4 and ND5 genes, while no amino acid difference was found in the ATP8 gene. Interestingly, the ATP8, which is generally considered one of the fastest evolving mtDNA genes in comparisons between distant species (Pesole et al., 1999; Saccone et al., 2000), showed little nucleotide divergence at intra- and inter-specific level in this study as well as two other fish species (Delarbre et al., 2001; Doiron et al., 2002) and humans (Mishmar et al., 2003), suggesting that this gene might not be informative at these or lower levels of divergence. On the other hand, the highly conserved cytochrome oxidase subunits 1, 2, and 3 (COI, COII, COIII) showed few of both nucleotide and amino acid changes. The analysis of nonsynonymous, (Ka) versus synonymous, (Ks) changes for all 13 mtDNA protein-coding genes revealed that the Ka/Ks ratio was slightly variable among genes, suggesting differential selective constraints (Fig. 3). However, the patterns of nucleotide substitution for each gene were similar among species pairs suggesting a constant mutational rate among these Gadidae species with the exception of the ND4L gene. For this gene, the high amino acid substitution relative to the nucleotide substitution may indicate some level of selective constraints. Different levels of selective pressures

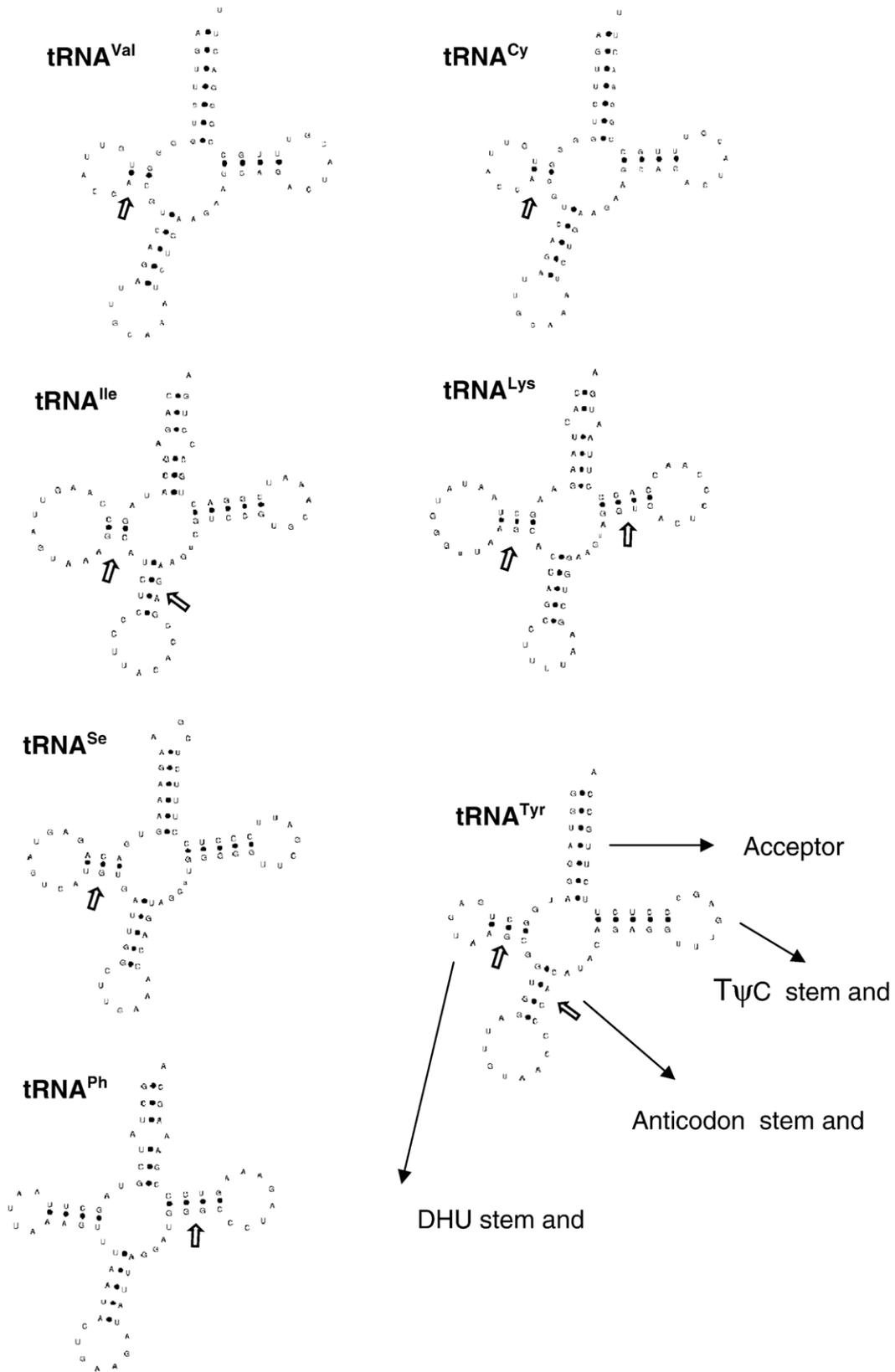


Fig. 1. Atypical secondary structures for seven inferred tRNAs of *Merlangius merlangius* and *Melanogrammus aeglefinus*. These cloverleaf structures are common to both gadidormes mitogenomes sequenced so far (*Gadus morhua*, Johansen and Bakke, 1996 and *Theragra chalcogramma*, Yanagimoto et al., 2004). Black points between nucleotide pairs indicate base pairing. Arrows indicate the size-reduced stems.

Table 3  
Tamura–Nei distances (d) and standard errors (between parenthesis), numbers of transitions (TS) and transversions (TV) among the four Gadidae species; *Gadus morhua* (Gm), *Theragra chalcogramma* (Tc), *Melanogrammus aeglefinus* (Ma) and *Merlangius merlangius* (Mm); for the complete mitochondrial (mt) genome (All), the combined transfer RNAs (tRNAs), the ribosomal genes (12S and 16S), the control region (CR) and the combined protein-coding genes (PCG)

		Gm_Tc	Ma_Mm	Tc_Mm	Gm_Mm	Tc_Ma	Ma_Gm
All (16773 bp)	d	0.043 (0.002)	0.062 (0.002)	0.0084 (0.002)	0.086 (0.002)	0.093 (0.002)	0.0094 (0.003)
	TS	566	756	977	993	1105	1100
	TV	117	213	304	323	311	328
tRNAs (1544 bp)	d	0.010 (0.003)	0.011 (0.003)	0.014 (0.003)	0.011 (0.003)	0.017 (0.003)	0.013 (0.003)
	TS	11	12	18	13	20	15
	TV	4	4	3	3	5	5
12S (953 bp)	d	0.005 (0.002)	0.015 (0.004)	0.014 (0.004)	0.013 (0.004)	0.014 (0.003)	0.013 (0.004)
	TS	5	13	11	9	11	10
	TV	0	1	3	3	2	2
16S (1677 bp)	d	0.006 (0.002)	0.028 (0.004)	0.014 (0.003)	0.015 (0.003)	0.022 (0.004)	0.022 (0.004)
	TS	7	37	18	19	27	28
	TV	3	8	5	6	9	8
CR (890 bp)	d	0.070 (0.009)	0.105 (0.011)	0.109 (0.012)	0.143 (0.014)	0.184 (0.016)	0.184 (0.016)
	TS	31	39	40	53	59	65
	TV	25	51	59	72	92	87
PCG (11472 bp)	d	0.059 (0.003)	0.086 (0.003)	0.117 (0.004)	0.123 (0.003)	0.136 (0.004)	0.138 (0.004)
	TS	509	660	853	876	971	972
	TV	92	170	212	232	218	240

The length of the alignment for each mtDNA portion is given in base pairs (bp).

acting on the mutational rate of the mtDNA regions have been previously reported in humans and *Drosophila* species (Ballard and Kreitman, 1995; Mishmar et al., 2003; Ruiz-Pesini et al., 2004).

The topology of the phylogenetic trees of the five species based on the entire mtDNA genomes, the combined PCG, the

12S and 16S ribosomal subunits and the combined tRNA genes is presented for the MP analysis, but was similar for all phylogenetic analyses (MP, NJ and ML) using either neighbor-joining (NJ) or maximum-likelihood (ML) analyses (Fig. 4). The entire mtDNA genomes and the combined PCG depicted two monophyletic clades, *Gadus* *Theragra* and

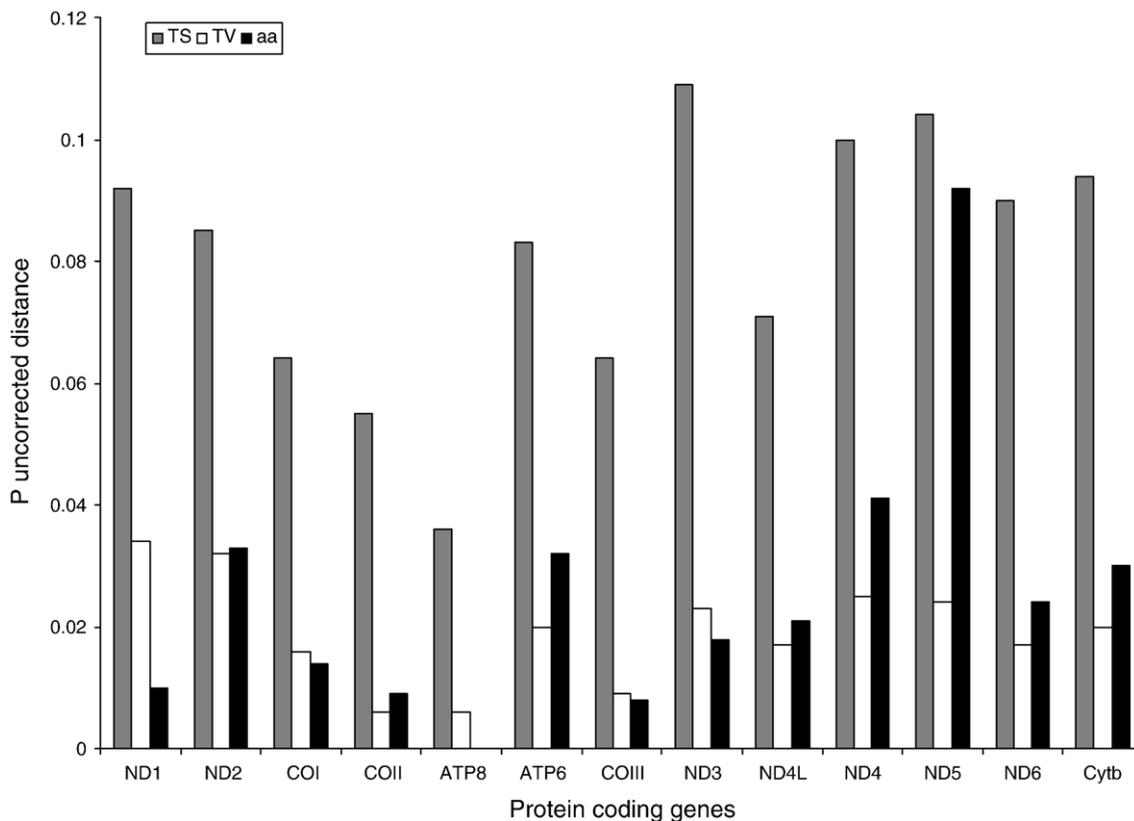


Fig. 2. Uncorrected p distances for nucleotide substitutions (transitions, TS and transversions, TV) and amino acid changes (aa) between the most divergent species pairs, *M. aeglefinus* and *G. morhua*, for each protein-coding gene. The genes appear in the order observed in the mitochondrial genome.

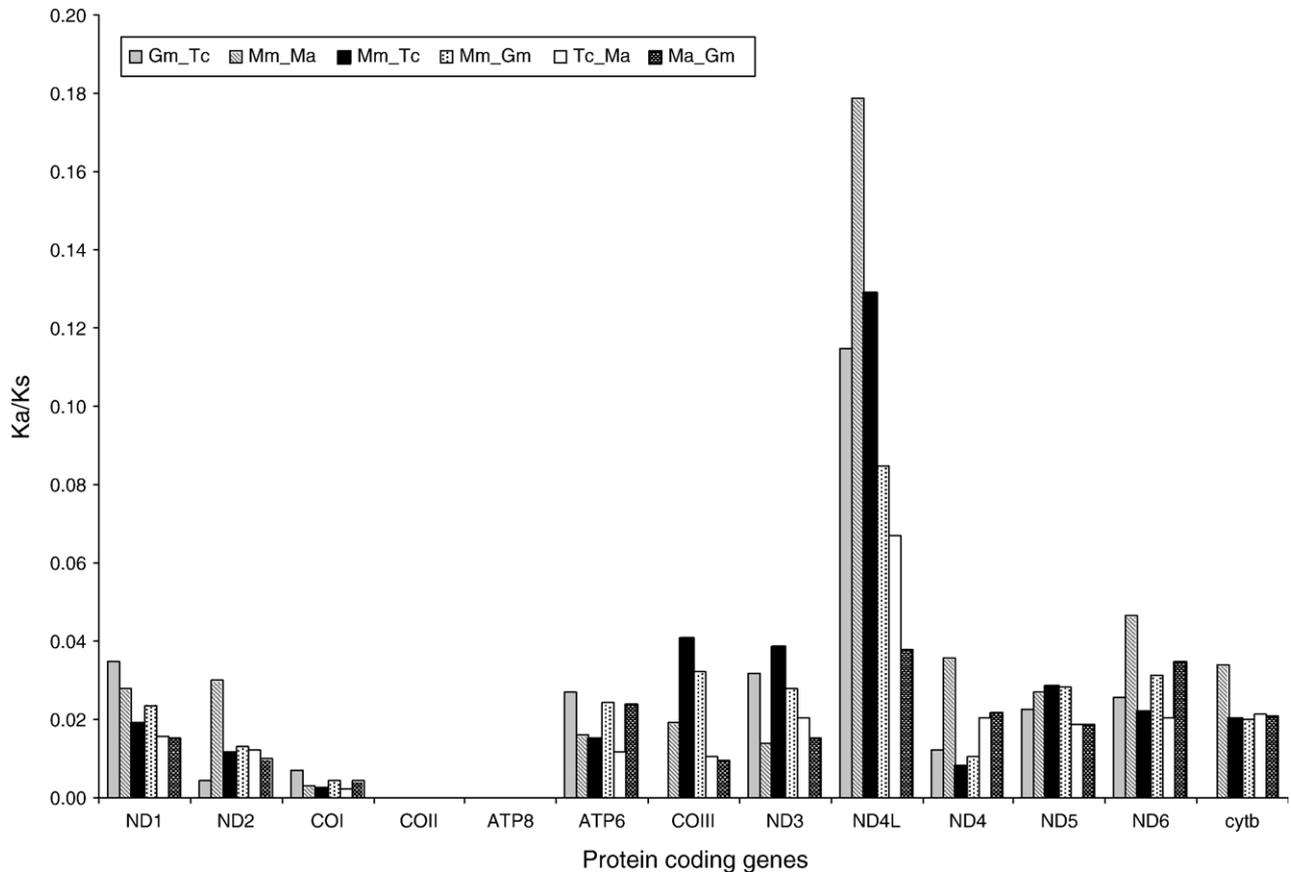


Fig. 3. Relationships between the number of amino acid differences per site (Ka) and the number of synonymous differences per site (Ks) of the protein-coding genes for the six pairs: *Gadus morhua* *Theragra chalcogramma* (Gm\_Tc); *Merlangius merlangus* *Melanogrammus aeglefinus* (Mm\_Ma); *T. chalcogramma* *M. merlangus* (Tc\_Mm); *G. morhua* *M. merlangus* (Gm\_Mm); *T. chalcogramma* *M. aeglefinus* (Tc\_Ma); *G. morhua* *M. aeglefinus* (Gm\_Ma). The genes appear in the order observed in the mitochondrial genome.

*Merlangius/Melanogrammus*. For these, identical topologies were recovered by all three phylogenetic methods, when all or only third positions were analysed, or by NJ when either only transversions or transitions were considered. Both clades were highly supported by both Bremer and bootstrap values, although the *G. morhua* *T. chalcogramma* clade was overall better supported (PBS = 142; Fig. 4). Furthermore, in the MP trees for each PCG (Fig. 4), most of the genes performed well in describing the phylogenetic relationships among the five species, as both clades observed matched those depicted by the combined PCG and the whole genomes. Similar MP phylogenies were found using either all codon positions or only the third one, except for the COIII gene that failed to group *Merlangius* and *Melanogrammus*, when only the third codon was used. When using NJ analyses based on both transitions and transversions, or on transitions alone, all topologies were similar to the MP method at all codon positions. On the other hand, when only transversions were used, the COI, ATP8, COIII, ND4L, ND5 and *cytb* failed to group *Merlangius* and *Melanogrammus*, suggesting that transversions may not be phylogenetic informative at this level of divergence. Using the ML analysis, all genes gave the same topology as MP and NJ analyses, except the ATP8 gene, that failed to group *Gadus* and *Theragra*.

For the analysis of the performance of each gene in depicting the expected phylogeny, the ND1, ND2 and ND4 genes were the best performers, with both the highest bootstraps and PBS values, while the ATP6, COII, ND6, 16S and tRNAs genes performed poorly, as they do not support the *Merlangius* *Melanogrammus* clade (negative PBS values). On the other hand, all genes support the *Gadus* *Theragra* clade (bootstrap values  $\geq 50\%$ ). Interestingly, [Zardoya and Meyer \(1996\)](#) found that ND2, ND4, ND5 were good phylogenetic performers in recovering both expected tetrapod and mammalian phylogenies, while ATP6 was a poor predictor. Also, in a previous analysis of the phylogenetic performance of mitochondrial genes for reconstructing higher-level relationships of teleosts, [Miya and Nishida \(2000\)](#) found that ND5 was the sole protein-coding gene that consistently reproduced the expected phylogeny, while ND6 and ATP6 occupied the less informative gene characterized by noticeably poor phylogenetic performances. The highly conserved ribosomal and transfer RNA genes performed badly and gave none or low support to the *Merlangius* *Melanogrammus* clade. While these genes are often used and proved powerful in higher-level phylogenetics ([Mindell and Honeycutt, 1990](#)), they may not be informative to detect the correct phylogenetic signal for the scale of evolutionary divergence analysed in this study ( $\sim 10$  my).

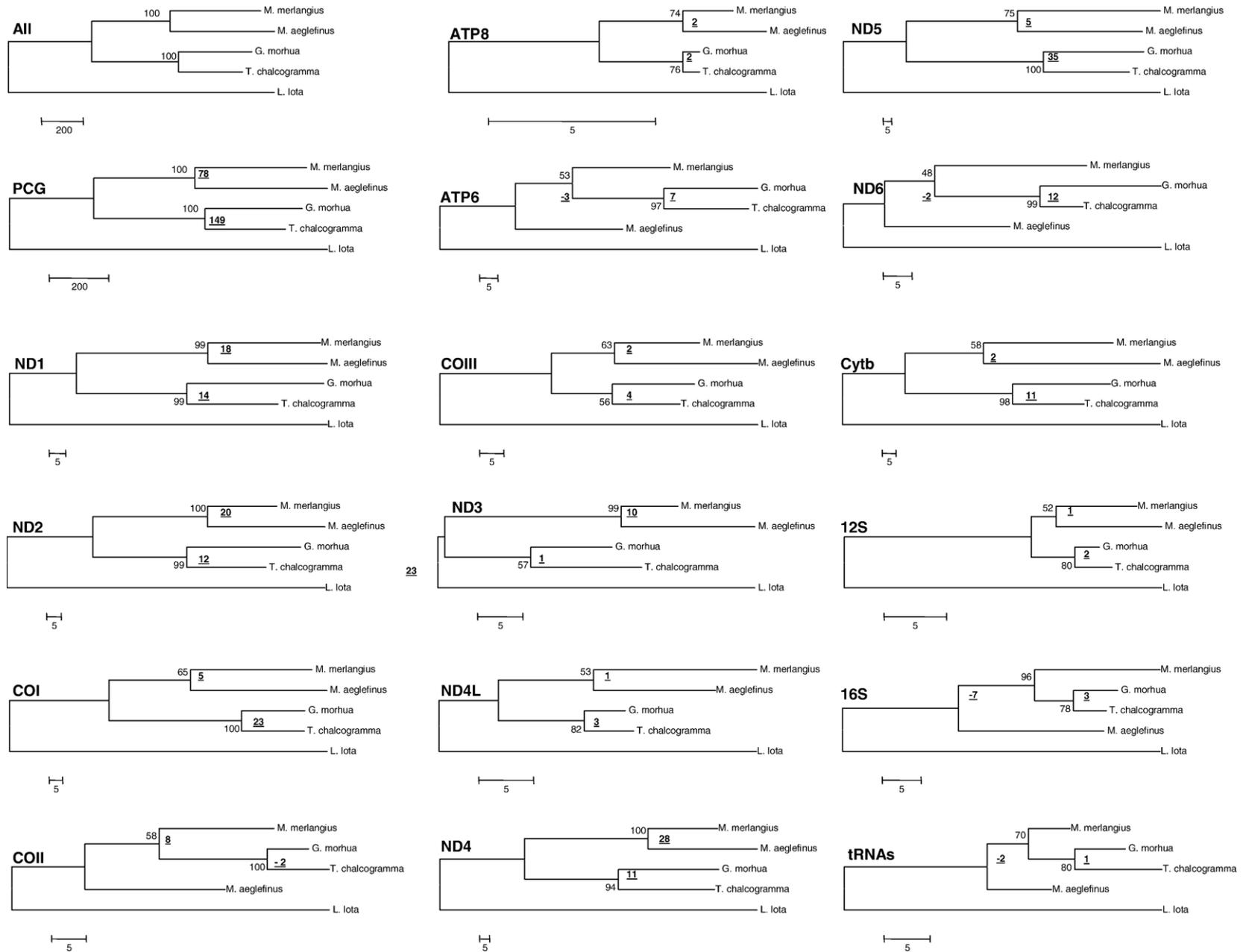


Fig. 4. Maximum-parsimony trees of four Gadidae species and one Lotidae outgroup taxon (*Lota lota*), for the whole mitochondrial genomes (All), for the combined protein-coding genes (PCG), for each of the 13 PCG, separately and for the ribosomal (12S and 16S) and transfer RNA (all combined) genes. Bootstrap values given above nodes were obtained with 1000 replicates. Underlined values within nodes indicated the partitioned Bremer support (PBS) at this node.

It has been suggested that the performance of a given protein-coding gene in recovering evolutionary relationships at different levels of divergences is also related to its length (Russo et al., 1996). This was partly confirmed by the fact that the “best” performing genes (ND1, ND2, ND4, COI and ND5) were also the longest ones, but, on the other hand, the less performing genes (ATP6, ND6 and COII) were not the shortest ones (see Table 1). We therefore decided to test whether the best performance of a gene was simply due to its larger number of informative sites. To this end, phylogenetic analyses were conducted again, using a subset of first 521 bp for each protein-coding gene. This length corresponds to the smallest, poorly performing gene (ND6). Interestingly, the shortened COI, cytb, ND5, COIII genes failed to group the *Merlangius\_Melanogrammus* clade, suggesting that using a short portion of a given mtDNA gene to infer phylogeny might be taken carefully and that such sequence may not be representative of the species tree. On the other hand, the best performers (ND1, ND2 and ND5) still give the expected phylogeny, pointing out that factors, other than the length may influence the performance of a gene. We also compared whether branch lengths may influence the performance of a given gene in depicting the true relationships. We found that the phylogenies obtained from the six best performing genes (ND1, ND2, COI, ND4, CYTB, ND5) were characterized by the longest branches, but the ATP6 gene with also large branching performed badly. In terms of levels of polymorphism for each PCG (Fig. 2) and the performance of a gene in reflecting the divergence between species pairs, we found that the most divergent genes (ND1, ND2, COI, ND4 and ND5) among the most divergent species pair comparison, *G. morhua*\_M. *aeglefinus* are also those that performed well in matching the phylogeny of the combined PCG. More specifically, the three best performers ND1, ND2 and ND4 (higher bootstrap and Bremer values) were those that showed the higher levels of transversions rates (Fig. 2). On the other hand, there was no consistent pattern among the other genes with variable mutational rates, with ND3 and COIII performing well, but COII, ATP6 and ND6 performing very poorly.

### 3.4. Conclusions

In this study, we report the whole mitochondrial genomes of two commercially important fish species, whiting and haddock, and depicted several evolutionary and structural patterns of their genes. The organization and gene content of the mtDNA genomes of *M. merlangius* and *M. aeglefinus* do not differ significantly from each other neither from those observed in the vast majority of fish species sequenced to date. By including previously published sequence data from the Atlantic cod (*G. morhua*) and walleye pollock (*T. chalcogramma*), we found that although the extent of sequence variation differed among the different protein-coding genes, the evolutionary rate of a given gene was relatively constant among all Gadidae species pairs examined. Finally, we observed that although most genes performed quite well in depicting the expected phylogeny and were quite variable in terms of length, substitution rate and branch sorting, the best performers were the largest genes and

were also characterized by a higher substitution rate (mostly transversions) and with trees characterized by longer external branches. These genes remained the best performers even when only the first 521 bp of the gene were used in the analysis. On the other hand, because some of the poorly performing genes (ATP6, ND6) could not be associated with any particular characteristic, we suggest that various factors (e.g. divergence level, lineage sorting and mutational constraints) may better explain their low efficiency in recovering phylogenetic relationships at the level of divergence considered in this study. Also, these results point out that the general value of analysing and comparing whole mitochondrial genomes at several levels of divergence.

### Acknowledgments

This research was a spin-off from DEFRA project M0146 Genetic identification of fish eggs by species-specific DNA markers for use in stock biomass assessments by the egg production method. We would especially like to thank the following people for their assistance in providing samples or access to sampling cruises: Phil Davison, Thomas Hurlbut, Martin Castonguay, Linda Despres, Anna Danielsdottir, Steine Tveite, Are Salthaug, Andrew Newton, Trevor Boon, Myrtle Boon, Richard Ayers, Audrey Geffen and Adam Corner.

### Appendix A

Primers' sequences (5'–3') and location

Primer names	Sequences (5'–3')	Region
GluDG	TGACTTGAARAACCAAYCGTTG	Cytb
CB3H	GGCAAATAGGAARTATCATTC	Cytb
L8 331	AAAGCRITYRGCCTTTTAAGC	ATPase
CO3.2	GTTAGTGGTCAKGGGCTTGGRCT	ATPase
H17	ACTTTCTAGGGTCCATC	tRNAPhe
L538	TTTTATACCCAAACCATCCGCCTG	12S
L19	CCACTAGCTCCCAAAGCTA	rRNA
12S760R	CAT CCC ATT CGC TAC ACC T	tRNAPhe
		12S
		rRNA
H1028	TAGCCAAGTTAGTAAAGCATCTC	12S
		rRNA
L2051	TGACTATAAGTTTAAACGGCCGCGG	16S
		rRNA
H2539	AGCCGCTATTAAGGGTTCGTTTGTTC AAC	16S
		rRNA
L3739	AACTTTTTACCTTTAACTTTGGCCC	ND1
H3884	GGCTCGAACCCATCCTCAGGAGATC	ND1
H4431	GCACCYTTTATCTTAAIRGTCAA	ND2
L5013	ATTTCTGCTATRCTTTTGCCCTG	ND2
H5144	ACCCACATCTTCTGTATGCAAAAC	tRNAAla
L5234	AACAGCTAAGCACTCAAACCAGCGA	tRNAAsn
H5765	CCYCCATCTTCTCTCTCCTTTTA	COI
L6763	GATATTCAGACTACCCGATGCCT	COI
H7026	CAAGCCAATCACATGACCACTCTGT	tRNASer
L7068	CACTCTGTCACTTTCTTCATGGGCT	tRNAAsp
L7963	ATGAGCAATTTTYCTAACTATTCTTCCC	ATPase
H9984	TCGGCTCAAGAGTCTGTGGTAAAT	tRNAArg
L11110	CTTTGAGGTGTAATTATAACRGGCT	ND4
ND5_11870F	AAG ACT GCT AAC TAT CAC CC	tRNALeu

(continued on next page)

## Appendix A (continued)

Primer names	Sequences (5'–3')	Region
H11899	GCAGCTATGCACCCATAACTTTAA	tRNA <sub>Leu</sub>
H12508	AGCCACTAAYATTAATAGCTGRGA	ND5
gapND5_12390F	TTGGTTGAGAGGGTGTGTYGGA	ND5
ND5_13310F	TTA GCC CCT TTT ACT TCC AC	ND5
gapND5_13270R	AAA AAG CGG GGT GTG CCT ATG	ND5
H13657	CACCCCTCTAGCCTCTGCAACAAG	ND5
L13697	TTAATTAAGACC–TACCTGACCTA	ND5
ND514130R	GAG GGA ATA TGA GTA CCA G	ND6
L14376	TAATGATGCAYTWGTTGATCTTCCC	Cytb
CytbF14960	CAGGCTCAAACAACCCSACA	Cytb
H15699	TACCCCTCTGAACATTAYTARTAG	tRNA <sub>Pro</sub>
HNEW1	TTTTAAAATATTRWTTATCGCKYGAGY	Control region
HNEW2	AAAATATTRWTTATCGCKYGAGY	Control region
LNEW1	CAAGAAATTGTAAACATAACCGGAC	Control region
LNEW2	ATRWATTWATAGAACAATATGGTT	Control region
DloopR15970	AGGACATAGATACATAAAGG	Control region
DloopR16170	GGATGGTGGGTCTTCTATGT	Control region
H16460	CGAYCTTCGGATTACAAGACCG	Control region
L16496	CGATGAGTTCCTAATATACCCCT	Control region
WhDL16670F	ATG AGT TCC TAA TAT ACC CCC	Control region
HdDL16680F	CCC CTG TTG ACT CTA ATT TG	Control region

## References

Arnason, E., 2004. Mitochondrial cytochrome b DNA variation in the high-fecundity Atlantic cod: transatlantic clines and shallow gene genealogy. *Genetics* 166, 1871–1885.

Arnason, E., Palsson, S., 1996. Mitochondrial cytochrome b DNA sequence variation of Atlantic cod, *Gadus morhua*, from Norway. *Mol. Ecol.* 5, 715–724.

Arnason, E., Petersen, P., Palsson, S., 1998. Mitochondrial cytochrome b DNA sequence variation of Atlantic cod, *Gadus morhua*, from the Baltic and the White seas. *Hereditas* 129, 37–43.

Arnason, E., Petersen, P., Kristinsson, K., Sigurgislaon, H., Palsson, S., 2000. Mitochondrial cytochrome b DNA sequence variation of Atlantic cod from Iceland and Greenland. *J. Fish Biol.* 56, 409–430.

Avise, J.C., 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. *Philos. Trans. R. Soc. Lond., B* 312, 325–342.

Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Syst. Biol.* 46, 654–673.

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, S., 2005. Molecular phylogenetics of Gadidae and related gadiformes based on mitochondrial DNA sequences. *Mar. Biotechnol.* 7, 61–69.

Bakke, I., Shields, G., Johansen, S., 1999. Sequence characterization of a unique intergenic spacer in gadiformes mitochondrial DNA. *Mar. Biotechnol.* 1, 411–415.

Ballard, J.W.O., Kreitman, M., 1995. Is mitochondrial DNA a strictly neutral marker? *TREE* 10, 485–488.

Bermingham, E., McCafferty, S., Martin, A., 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian isthmus. In: Kocher,

Thomas D., Stepien, Carol A. (Eds.), *Molecular Systematics of Fishes*. Academic Press, San Diego, pp. 113–128.

Cao, Y., Adachi, J., Janke, A., Paabo, S., Hasegawa, M., 1994. Phylogenetic relationships among Eutherian order estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Mol. Evol.* 39, 519–527.

Carr, S., Marshall, H., 1991. A direct approach to the measurement of genetic variation in fish populations: applications of the polymerase chain reaction to study of Atlantic cod, *Gadus morhua* L. *J. Fish Biol.* 39, 101–107.

Carr, S., Snellen, A., Howse, K., Wroblewski, J., 1995. Mitochondrial DNA sequence variation and genetic stock structure of Atlantic cod (*Gadus morhua*) from Bay and offshore locations on the Newfoundland continental shelf. *Mol. Ecol.* 4, 79–88.

Carr, S., Kivlichan, D., Pepin, P., Crutcher, D., 1999. Molecular systematics of gadid fishes: implications for the biogeographic origins of Pacific species. *Can. J. Zool.* 77, 19–26.

Cohen, D.M., Inada, T., Iwamoto, T., Scialabba, N., 1990. FAO species catalogue. Vol. 10. Gadiform fishes of the world (Order Gadiformes). An annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to date. FAO Fish. Synop. 10 (125) 442 pp.

Cummings, M., Otto, S., Wakeley, J., 1995. Sampling properties of DNA sequence data in phylogenetic analysis. *Mol. Biol. Evol.* 12, 814–822.

Delarbre, C., Rasmussen, A., Arnason, U., Gachelin, G., 2001. The complete mitochondrial DNA genome of the hagfish *Myxine glutinosa*: unique features of the control region. *J. Mol. Evol.* 53, 634–641.

Doiron, S., Bernatchez, L., Blier, P., 2002. A comparative genomics analysis of the potential adaptive value of Arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill). *Mol. Biol. Evol.* 19, 1902–1909.

Fox, C.J., Taylor, M.I., Pereyra, R., Villasana-Ortiz, M.I., Rico, C., 2005. TaqMan DNA technology confirms likely over-estimation of cod (*Gadus morhua* L.) egg abundance in the Irish Sea: implications for the assessment of the cod stock and mapping of spawning areas using egg based methods. *Mol. Ecol.* 14, 879–884.

Graur, D., Li, W.H., 1999. *Fundamentals of Molecular Evolution*. Sinauer Associates, Sunderland, Mass.

Inoue, J., Miya, M., Tsukamoto, K., Nishida, M., 2000. Complete mitochondrial DNA sequence of the Japanese sardine, *Sardinops melanostictus*. *Fish. Sci.* 66, 924–932.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2001. Complete mitochondrial DNA sequence of Conger myriaster (Teleostei: Anguilliformes): novel gene order for vertebrate mitochondrial genomes and the phylogenetic implications for anguilliform families. *J. Mol. Evol.* 52, 311–320.

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.

Kim, I., et al., 2004. The complete mitochondrial genome of the javeline goby, *Acanthogobius hasta* (Perciformes, Gobiidae) and phylogenetic considerations. *Gene* 336, 147–153.

Kumar, S., Tamura, K., Jakobsen, I., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17, 1244–1245.

Lee, W., Conroy, J., Howell, W., Kocher, T., 1995. Structure and evolution of teleost mitochondrial control regions. *J. Mol. Evol.* 41, 54–66.

Lee, J., et al., 2001. The complete DNA sequence of the mitochondrial genome of the self-fertilizing fish, *Rivulus marmoratus* (Cyprinodontiformes, Rivulidae) and the first description of duplication of a control region in fish. *Gene* 280, 1–7.

Lowe, T., Eddy, S., 1997. tRNA<sub>Scan</sub>-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964.

Manchado, M., Catanese, G., Infante, C., 2004. Complete mitochondrial DNA sequence of the Atlantic bluefin tuna, *Thunnus thynnus*. *Fish. Sci.* 70, 68–73.

McKnight, M.L., Shaffer, H., 1997. Large, rapidly evolving intergenic spacers in the mitochondrial DNA of the salamander family Ambystomatidae (Amphibia: Caudata). *Mol. Biol. Evol.* 14, 1167–1176.

Meyer, A., 1994. Shortcomings of the cytochrome b as a molecular marker. *Trends Ecol. Evol.* 9, 278–280.

Mindell, D.P., Honeycutt, R.L., 1990. Ribosomal RNA in vertebrates: evolution and phylogenetic applications. *Ann. Rev. Ecol. Syst.* 21, 541–566.

Mishmar, D., et al., 2003. Natural selection shaped regional mtDNA variation in humans. *Proc. Natl. Acad. Sci. U. S. A.* 100, 171–176.

- Miya, M., Nishida, M., 1999. Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. *Mar. Biotechnol.* 1, 416–426.
- Miya, M., Nishida, M., 2000. Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. *Mol. Phylogenet. Evol.* 17, 437–455.
- Miya, M., Kawaguchi, A., Nishida, M., 2001. Mitogenomic exploration of higher teleostean phylogeny: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. *Mol. Biol. Evol.* 18, 1993–2009.
- Miya, M., et al., 2003. Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 26, 121–138.
- Moller, P., Jordan, A.P., Steffensen, J., 2002. Phylogenetic position of the cryopelagic codfish genus *Arctogadus* Drjagin, 1932 based on partial mitochondrial cytochrome b sequences. *Polar Biol.* 25, 342–349.
- Nolf, D., Steurbaut, E., 1989. Importance and restriction of the otolith-based fossil record of gadiform and ophidiiform fishes. In: Cohen, D. (Ed.), *Papers on the Systematics of Gadiform Fishes*. Science Series of the Natural History Museum Los Angeles County, vol. 32, pp. 47–58.
- Pesole, G., Gissi, C., De Chirico, A., Saccone, C., 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. *J. Mol. Evol.* 48, 427–434.
- Pollock, D.D., Eisen, J.A., Doggett, N.A., Cummings, M.P., 2000. A case for evolutionary genomics and the comprehensive examination of sequence biodiversity. *Mol. Biol. Evol.* 17, 1776–1788.
- Posada, D., Crandall, K., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Quinn, T., 1997. Molecular evolution of the mitochondrial genome. In: Mindell, D.F. (Ed.), *Avian Molecular Evolution and Systematics*. Academic Press, pp. 3–28.
- Rico, C., Kuhnlein, U., FitzGerald, G.J., 1992. Male reproductive tactics in the threespine stickleback: an evaluation by DNA fingerprinting. *Mol. Ecol.* 1, 79–87.
- Rozas, J., Sanchez-Delbarrio, J., Messeguer, X., Rozas, R., 2003. DNASP, DNA polymorphism analysis by the coalescent and other methods. *Bioinformatics* 19, 2496–2497.
- Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., Wallace, D.C., 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 303, 223–226.
- Russo, C., Takezaki, N., Nei, M., 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. *Mol. Biol. Evol.* 13, 525–536.
- Saccone, C., Gissi, C., Lanave, C., Larizza, A., Pesole, G., Reyes, A., 2000. Evolution of the mitochondrial genetic system: an overview. *Gene* 261, 153–159.
- Sigurgislaon, H., Arnason, E., 2003. Extent of mitochondrial DNA sequence variation in Atlantic cod from the Faroe Islands: a resolution of gene genealogy. *Heredity* 91, 557–564.
- Sorenson, M.D., 1999. *TreeRot*. Version 2. Boston University, Boston, Mass.
- Swofford, D.L., 1999. *PAUP\* 4.0*. Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer, Sunderland, Mass.
- Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.
- Taylor, M.I., Fox, C.J., Rico, I., Rico, C., 2002. Development of species specific Taqman probes for (*Gadus morhua* L.), haddock (*Melanogrammus aeglefinus* L.) and whiting (*Merlangius merlangus* L.). *Mol. Ecol. Notes* 2, 599–601.
- Wheeler, A.C., 1978. *Key to the Fishes of Northern Europe: A Guide to the Identification of More than 350 Species*. F. Warne, London, ISBN 0723220972, 0723220646. 380pp.
- 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., Meyer, A., 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Mol. Biol. Evol.* 13, 933–942.
- Zardoya, R., Malaga-Trillo, E., Veith, M., Meyer, A., 2003. Complete nucleotide sequence of the mitochondrial genome of a salamander, *Mertensiella luschani*. *Gene* 317, 17–27.
- Zwanenburg, K., Bentzen, P., Wright, J., 1992. Mitochondrial DNA differentiation in Western North Atlantic populations of haddock (*Melanogrammus aeglefinus*). *Can. J. Fish. Aquat. Sci.* 49, 2527–2537.