Molecular Phylogenetics and Evolution 95 (2016) 161–170

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the NADH2 gene in whitefish (*Coregonus* ssp.) ☆

Magnus W. Jacobsen [a,*], Rute R. da Fonseca [b], Louis Bernatchez [c], Michael M. Hansen [a]

[a] Department of Bioscience, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus C, Denmark
[b] Department of Bioinformatics and RNA Biology, University of Copenhagen, Ole Maaløesvej 5, 2200 København N, Denmark
[c] Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Pavillon Charles-Eugène-Marchand, 1030, Avenue de la Médecine, Université Laval, Québec, Québec, Canada G1V 0A6

ARTICLE INFO

Article history:
Received 18 August 2015
Revised 6 November 2015
Accepted 13 November 2015
Available online 2 December 2015

Keywords:
Mitogenome
Positive selection
Relaxed purifying selection
dN/*dS*
ND2
Coregonus ssp.

ABSTRACT

Several studies have recently reported evidence for positive selection acting on the mitochondrial genome (mitogenome), emphasizing its potential role in adaptive divergence and speciation. In this study we searched 107 full mitogenomes of recently diverged species and lineages of whitefish (*Coregonus* ssp.) for signals of positive selection. These salmonids show several distinct morphological and ecological differences that may be associated with energetics and therefore potentially positive selection at the mitogenome level. We found that purifying selection and genetic drift were the predominant evolutionary forces acting on the analyzed mitogenomes. However, the NADH dehydrogenase 2 gene (ND2) showed a highly elevated *dN*/*dS* ratio compared to the other mitochondrial genes, which was significantly higher in whitefish compared to other salmonids. We therefore further examined nonsynonymous evolution in ND2 by (i) mapping amino acid changes to a protein model structure which showed that they were located away from key functional residues of the protein, (ii) locating them in the sequences of other species of fish (Salmonidae, Anguillidae, Scombridae and Percidae) only to find pronounced overlap of nonsynonymous regions. We thus conclude that relaxed purifying selection is driving the evolution of ND2 by affecting mostly regions that have lower functional relevance.

1. Introduction

The vertebrate mitochondrial genome (mitogenome) consists of 37 genes (Ballard and Whitlock, 2004). Of these, 13 encode subunits of protein complexes directly involved in the oxidative phosphorylation (OXPHOS) pathway responsible for most of the energy (ATP) produced in the eukaryotic cell (Saraste, 1999). Despite the important functions of its genes, the mitogenome has historically been assumed to evolve neutrally (Ballard and Whitlock, 2004; Galtier et al., 2009). However, several studies have shown discrepancies from solely neutral evolution (Ballard and Whitlock, 2004; Galtier et al., 2009). Analyses of *dN*/*dS* between the mitochondrial coding genes normally demonstrate values $\ll 1$ in e.g. fishes and mammals (e.g. da Fonseca et al., 2008; Sun et al., 2011), which is evidence for strong purifying (negative) selection. This is in line with pathophysiology research showing associations between amino acid changes in mitochondrial coding genes and disease (Jonckheere et al., 2012; Smeitink et al., 2001). Given strong evolutionary constraints, most amino acid (nonsynonymous) changes are likely to be nearly neutral; i.e. slightly deleterious or situated in regions of less overall importance for protein function.

Although purifying selection may be the predominant form of selection acting at the mitogenome level it does not exclude the possibility for positive selection acting on individual genes or single codon positions (Ballard and Whitlock, 2004; Galtier et al., 2009). However, the mitogenome has an effective size (N_e) four times smaller compared to a diploid nuclear genome, due to a haploid nature and maternal inheritance (Ballard and Whitlock, 2004; Birky et al., 1983; Galtier et al., 2009). Thus, compared to its nuclear counterparts, mitochondrial genes are likely to be more influenced by genetic drift than selection. As a consequence some studies have argued that positive selection should be more common in organisms showing high N_e, like many invertebrate species. However, empirical studies find only limited support for this (Bazin et al., 2006; Meiklejohn et al., 2007; Piganeau and Eyre-Walker, 2009). Using different combinations of analyses, several

☆ This paper was edited by the Associate Editor Giacomo Bernardi.

* Corresponding author.
E-mail address: magnus.jacobsen@bios.au.dk (M.W. Jacobsen).

http://dx.doi.org/10.1016/j.ympev.2015.11.008
1055-7903/

studies have reported evidence for positive selection in the coding genes of the mitogenome in vertebrate species with presumably lower effective population sizes. As such, evidence for positive selection has been found in mammals (da Fonseca et al., 2008; Fontanillas et al., 2005; Foote et al., 2011), and fishes (Belanger-Deschenes et al., 2013; Dalziel et al., 2006; Gagnaire et al., 2012; Garvin et al., 2011; Jacobsen et al., 2014; Jacobsen et al., 2015) probably relating to thermal adaption, or aerobic requirements linked to migration behavior, swimming speed, size and diet. Also, given the tight link between mitochondrial and nuclear genes in the oxidative phosphorylation complex, possibilities for co-evolution exist (Ballard and Whitlock, 2004; Gagnaire et al., 2012; Galtier et al., 2009; Rand et al., 2004; Teacher et al., 2012), and positive selection or fixation of deleterious mutations in either genome can lead to subsequent selection acting on the other (Rand et al., 2004).

In this study we investigate the evolutionary patterns at the mitogenome level, across different species and ecotypes of whitefish (*Coregonus* ssp.). Whitefish are salmonids, show enormous phenotypic variation and are important models for studying the genetics of speciation (e.g. Gagnaire et al., 2013; Jacobsen et al., 2012; Østbye et al., 2006; Østbye et al., 2005; Pigeon et al., 1997; Rogers and Bernatchez, 2005, 2007). Several distinct ecotypes (or species) are known to occur both allo- and sympatrically. Among these are the 'Atlantic' and the 'Acadian' mtDNA lineages of the American lake whitefish (abbreviated ALW) (*Coregonus clupeaformis*) (Bernatchez and Dodson, 1990, 1994) that evolved in allopatry during the last Glaciation >50 Kya (Jacobsen et al., 2012). Both lineages can exhibit dwarf-sized pelagic or larger sized benthic forms, however, when occurring in sympatry individuals belonging to the Acadian lineage tend to have evolved toward the dwarf phenotype (Bernatchez and Dodson, 1990; Lu et al., 2001). Also in Europe several differentiated forms exist (Himberg and Lehtonen, 1995; Kottelat and Freyhof, 2007; Østbye et al., 2006; Østbye et al., 2005) like the North Sea houting (abbreviated NSH) (*Coregonus oxyrhynchus*, sometimes referred to as *C. oxyrhinchus*), a recently diverged species or ecotype of the European lake whitefish (abbreviated ELW) (*Coregonus lavaretus*) (Hansen et al., 2008; Jacobsen et al., 2012).

Both European and American whitefish forms show several phenotypic differences likely linked to metabolism, which suggests the possibility for positive selection to be acting at the mitogenome level. ALW shows differences in both age at maturity and growth rate (Trudel et al., 2001), with larger benthic forms maturing later and growing slower. In Europe, the NSH shows older age at maturation compared to Danish ELW (Hansen et al., 2008; Hansen et al., 2006). Moreover, as the only representative within the ELW-complex it migrates into high salinity seawater (>33‰) (Jensen et al., 2003) which could affect energy requirements related to osmoregulation and migration over longer distances in the marine environment.

Phenotypic differences in whitefish have in several cases been shown to have a genetic and possibly adaptive basis. QTL studies show significant association between microsatellite markers and e.g. growth rate, swimming behavior (habitat selection and predator avoidance), and life history (onset of maturity and fecundity) in sympatric 'Atlantic' and 'Acadian' lineages of ALW (Rogers and Bernatchez, 2007). Moreover, differences in gene expression have been reported between different whitefish species and ecotypes (Derome et al., 2006; Evans and Bernatchez, 2012; Jensen et al., 2013). Many of these genes relate to metabolism and likely reflect selection for optimal energy production. The genes include several involved in OXPHOS like the mitochondrial NADH dehydrogenase 5 (ND5, complex I), Cytochrome b (CYTB, complex III) and ATPase 8 genes (ATP8, complex V), as well as the nuclear Cytochrome b-c1 subunit 8 and Cytochrome c oxidase subunit 6A (complex IV) that all are upregulated in dwarf compared to normal ALW (Evans and Bernatchez, 2012). Also ATP synthase subunits g and c (complex V) show divergent expression patterns between Danish NSH vs. ELW (Jensen et al., 2013) and dwarf vs. normal ALW (Derome et al., 2006). However, although adaptive divergence as such seems evident at the regulatory level, no studies have so far investigated the possibility for positive selection at the functional level within the 13 mitochondrial genes.

Here we investigate potential selection at the mitogenome level in different species and ecotypes of whitefish by analyzing 107 mitogenome sequences. The dataset includes several distinct species and ecotypes that have diverged at different timescales; from few thousands to more than one million years ago. As such, the data constitute an interesting resource for elucidating evolution and selection at the mitogenome level over time. Given the phenotypic differences related to metabolism we especially wanted to test whether positive selection had occurred at the coding gene level. First, we tested the overall hypothesis of neutral evolution by comparing the number of mutations (total and synonymous) with the length of the respective genes, as well as the number of nonsynonymous change compared to the synonymous changes. Second, we analyzed the direction of amino acid changes at the phylogenetic level. We conducted codon-based analyses of diversifying and episodic positive selection, in order to investigate if positive selection is acting in the coding genes. Finally, as the NADH dehydrogenase 2 gene (ND2) showed exceptionally high dN/dS we further evaluated the amino acid changes within this gene by modeling the amino acid changes within the protein structure. In order to further investigate whether these changes were adaptive, we compared the observed patterns within the whitefishes with the overall pattern across 15 salmonid species, as well as three other families of fish (Scombridae, Percidae and Anguillidae). In this comparison, overlap of nonsynonymous changes between taxa was used as a proxy for regions likely under less conservative evolutionary constraints.

2. Material and methods

2.1. Sampling and sequences

The main dataset consisted of 107 mitogenome sequences of whitefish (*Coregonus* sp.) (Genbank accession No. NC_002646.1 and JQ661382–JQ661487) (Jacobsen et al., 2012; Miya and Nishida, 2000) and have previously been used to analyze the evolutionary history between European and American whitefish (Jacobsen et al., 2012). These sequences represent a total of 12 populations among which 6 populations (84 sequences) were derived from the Jutland Peninsula, Denmark. Among the sequences, 21 represent NSH (*C. oxyrhynchus*) and 73 represent ELW (*C. lavaretus*) from Denmark. A further 16 sequences of ELW are from the Baltic Sea region, one is from a lake in the Czech Republic (Miya and Nishida, 2000) and finally 6 sequences are from the closely related ALW (*C. clupeaformis*) with 3 individuals representing dwarf Acadian forms from Témiscouata lake and 3 representing the normal form belonging to the Atlantic lineage from Aylmer lake and Ross lake (Jacobsen et al., 2012) (for a overview over the phylogenetic relationships see Fig. 2). In order to compare the patterns of sequence evolution in whitefish, 15 salmonid mitogenome sequences were downloaded from Genbank and analyzed. These sequences belonged to different species or subspecies from four different salmonid genera; *Oncorhynchus* (9 species), *Salvelinus* (2 species), *Salmo* (2 species) and *Thymallus* (2 species). One sequence of northern pike (*Esox lucius*) were also downloaded and used as outgroup. Furthermore, ND2 gene sequences were also downloaded from Genbank from 18 species of freshwater eels

(Anguillidae), 24 species of tuna and mackerels (Scombridae) and 46 species of perciform fishes (Percidae) for downstream analysis (see Tables S1A–D for species and accession numbers for all downloaded sequences).

2.2. Rates of evolution

To compare the evolutionary rates between genes, the entire whitefish dataset ($N = 107$) was split into each of 13 coding genes; the two ribosomal RNAs; the control region and all the combined tRNAs using GENEIOUS PRO 5.4.6 (BioMatters, 2012). Measurements of the number of haplotypes (Nh), number of segregating sites (S), theta (θ) based on the finite sites model (Tajima, 1996) and Tajima's D (Tajima, 1989) were all calculated using the programme DnaSP v5.1 (Librado and Rozas, 2009). Also the numbers of nonsynonymous and synonymous substitutions for the coding genes were calculated in DnaSP v5.1 and average pairwise dN/dS calculated between the different haplotypes for each gene. As the same individuals were used, and as recombination is almost non-existent in the mitochondrial DNA, N_e ($N_e = 4\theta\mu$) should be similar for the individual genes. Thus, the θ-values were used to infer the relative mutation rates (per-generation mutation rates) of the individual genes relative to the whole mitogenome. This was conducted using the equation $\mu_{gene} = ((\mu_{mitogenome} \times \theta_{mitogenome})/\theta_{gene})$.

Using data from all 13 coding genes, neutral evolution in whitefish was assessed by conducting either linear regression or Spearman's rank analysis (when normality of the data was not met) in PAST (Hammer et al., 2001) between: (1) the number of mutations vs. length in bases, (2) the number of synonymous mutations vs. the length in bases and (3) the number of nonsynonymous changes vs. synonymous changes.

To compare the pattern to sequence evolution between other salmonids, linear regression and dN/dS analyses as described above were also conducted using the 15 salmonid sequences.

2.3. Direction of nonsynonymous change

All amino acid changes were identified using DnaSP version 5.1. Using the different whitefish mitogenome haplotypes ($N = 53$), a maximum likelihood phylogeny was generated in the program MEGA 5 (Tamura et al., 2011). The analysis was based on the TrN + G substitution model with eight gamma categories (Jacobsen et al., 2012). Subsequently the ancestral sequence of all the whitefish sequences was analyzed using Ancestral Sequence Reconstruction (ASR) software in HyPHY (Pond and Frost, 2005) implemented in the DataMonkey server (http://www.datamonkey.org/dataupload.php). The closely related European and Arctic grayling *Thymallus thymallus* and *Thymallus arcticus* sequences were used as outgroups (Genbank accession No. NC_012928 and NC_012929) (Yasuike et al., 2010). The ancestral sequence was then used to infer the direction of amino acid changes between the European and the North American sequences.

2.4. Analyses of positive selection

To analyze the 13 coding genes of from the whitefish dataset for positive selection, three codon-based selection tests were applied to the data: TreeSAAP version 3.2 (Woolley et al., 2003), FUBAR (Murrell et al., 2013) and MEME (Murrell et al., 2012). A phylogenetic tree identical to that in Jacobsen et al. (2012) was used in all analyses and only different haplotypes were used.

TreeSAAP (Woolley et al., 2003) relies on the MM01 model implemented in BASEML (part of the PAML package; Yang, 1997) and uses a phylogeny to reconstruct the most likely ancestral states for the gene sequences under investigation. It then assigns weight values to the nonsynonymous codon changes, for which overall physicochemical effects are assessed using a model with 31 physicochemical amino acid properties. The overall amino acid replacements are given a score from 1 to 8, with eight as the most significant change. Significant deviation from neutral evolution is tested via a z-score and interpreted as a result of positive selection. To ensure conservative calling of positive selected codon sites only amino acid changes with a score between 6 and 8 and with a positive z-score < 0.001 was used (Woolley et al., 2003). Due to the shallow divergence in the dataset, the phylogeny based on the whole mitogenome sequences was used to infer the evolution of the samples. Analyses were conducted for all the 13 coding genes.

FUBAR (Fast Unbiased Bayesian AppRoximation) is a codon-based maximum likelihood method that allows dN/dS (ω) to vary over each codon across a gene according to a number of predefined site classes given *a priori*. This allows for testing codons for positive ($dN/dS > 1$) or purifying selection ($dN/dS < 1$) (Murrell et al., 2013). While FUBAR posits that e.g. positive selection remains constant throughout time (affects most lineages in a phylogenetic tree) the MEME (Mixed Effects Model of Evolution) test allows the distribution of ω to vary over sites and moreover from branch to branch, which makes it possible to detect episodic selection (Murrell et al., 2012). Both FUBAR and MEME were analyzed using the HyPHY package as implemented on the DataMonkey server (http://www.datamonkey.org/dataupload.php). Best-fit substitution models for each dataset were also assessed here and used in the analyses. Significance was assessed by posterior probability > 0.9 (FUBAR) and P-value < 0.05 (MEME).

The minimum age of each candidate amino acid change, corresponding to the TMRCA of the specific clade from which individuals shared the same nonsynonymous substitution, was chosen based on estimates by Jacobsen et al. (2012) using a mutation rate of $1 * 10^{-8}$ sub/site/year.

2.5. Analysis of selection at the ND2 gene

As the only possible outlier in terms of dN/dS in whitefish, the NADH dehydrogenase 2 gene (ND2) was examined in more detail. In order to test whether the ND2 gene showed a significantly higher dN/dS in whitefish compared to other salmonids we used the program CODEML in PAML (Yang, 1997). We used a topology identical to Yasuike et al. (2010) but grouped whitefish as a sister-species to graylings following Campbell et al. (2013) that used full mitochondrial genomes to evaluate the phylogenetic relationships of Esociformes and Salmoniformes. A total of 6 whitefish sequences were used; one from each of the Acadian and Atlantic lineages and one from each of the two Danish clades and Baltic Sea clades. None of the sequences showed any unshared nonsynonymous changes. Northern pike was used as an out-group. We used likelihood ratio tests (LRT) to compare and test significance between two different models of dN/dS: (1) one dN/dS ratio across all branches (*Single ratio*), (2) one rate for all branches within the whitefish and one for all other branches (*Whitefish vs. other*). As higher dN/dS could be a result of higher genetic drift we conducted a control by analyzing the concatenated alignment constituting all other 12 mitochondrial genes (excluding ND2) and conducted LRTs using the same models. The ND2 dataset was also analyzed in HyPhy using the MEME and FUBAR test in order to test for episodic and diversifying positive selection across the entire salmonid phylogeny.

Furthermore, the codons subjected to amino acid changes in whitefish were compared to the sites showing change across the salmonid alignment. To test if replacements occurred randomly compared to the other salmonid species a chi square test (χ^2) were performed between the numbers of shared codon positions showing nonsynonymous changes, compared to the numbers of shared changes expected by chance alone (frequency of codon showing

nonsynonymous sites). The same test was also conducted for the three other alignments of freshwater eels (family; Anguillidae), tuna and mackerels (family; Scombridae) and perciform fishes (family; Percidae). The number of nonsynonymous changes occurring over each codon position in each of the four alignments (Salmonidae, Anguillidae, Scombridae and Percidae) was assessed using HyPhy, as implemented in MEGA 5, to account for unobserved changes. All alignments were aligned to the whitefish ND2 gene using GENEIOUS PRO 5.4.6.

A model of the ND2 subunit was built with SwissModel (Schwede et al., 2003) using the D chain of the PDB entry 3RKO corresponding to NADH-Quinone oxidoreductase subunit N (NuoN) (Efremov and Sazanov, 2011). Subsequently the amino acid residues fixed in whitefish were compared to the residues suggested to participate in proton translocation (E133, K217, K247 and K395 in NuoN).

3. Results

3.1. Rates of evolution

The relative evolutionary rate ($\mu_{Relative}$) differed across regions (Table 1). Overall, ND genes, CYT B and the control region evolved faster than the other regions, with ND1 and ND2 evolving the fastest, even faster than the control region. COX genes and tRNAs evolved ca. 2–3 times slower than the ND genes. The slowest evolving regions were the ribosomal genes and the ATP8 gene (Table 1). No nonsynonymous substitutions were observed in the ATP8 and COX1 genes. Tajimas *D* revealed little evidence for deviations from neutrality for all examined regions, with 12S rRNA as the only exception ($P < 0.05$) (Table 1). The number of nonsynonymous changes was highest for the ND2 gene, which showed a higher average pairwise *dN*/*dS* than for the other genes (0.352 vs. 0–0.139, Table 1). This was also found in the salmonid dataset, although less pronounced (0.052 vs. 0.039–0.008, Table 1).

Normality was met for gene length, total and synonymous mutations but not for the nonsynonymous mutations. Thus, the correlation between synonymous and nonsynonymous mutations was analyzed via Spearman rank analysis. Both the overall number of mutations (Fig. 1(A)) and the number of synonymous changes (Fig. 1(B)) showed significant linear relationships with the length of the genes. The number of nonsynonymous changes was also significantly correlated with the number of synonymous changes (Fig. 1(C)) although not as pronounced. In this analysis ND2 showed an exceptionally high relative proportion of nonsynonymous changes compared to the other genes (Fig. 1(C), Table 1). The salmonid dataset showed the same overall patterns (Fig. S1).

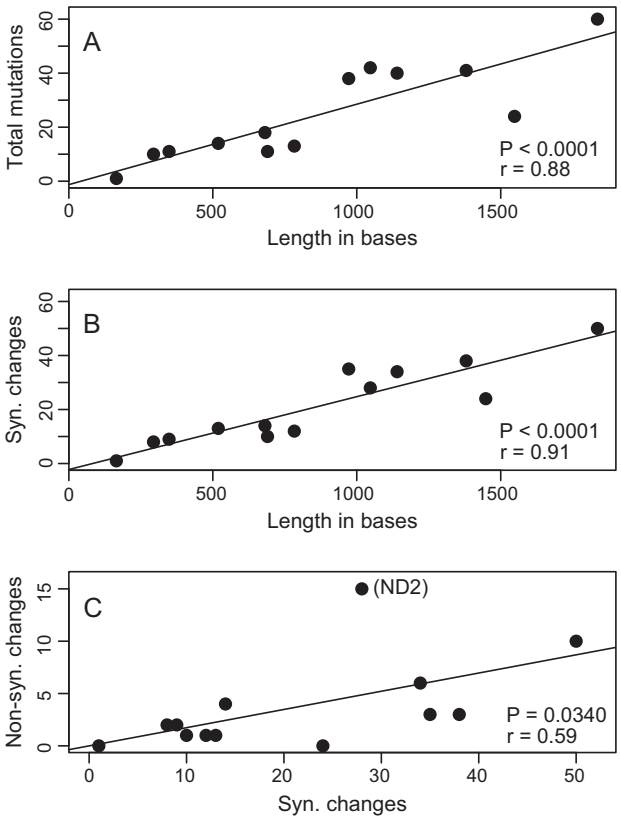

Fig. 1. Correlation analyses. The dots denote the individual values for the 13 different mitochondrial genes in whitefish and the line the best-fitted line. (A) Correlation between total number of mutations and length in bases of the genes. (B) Correlation between synonymous mutations and length in bases of the genes. (C) Correlation between nonsynonymous and synonymous mutations. The correlation coefficient '*r*' and *P*-values are shown in the lower right corner for linear regression (A + B) and Spearman rank test (C). The position of the ND2 gene is shown.

Table 1
Summary statistics for the whitefish mitogenome sequences and regions. The *dN*/*dS* is the average values of all pairwise comparisons.

Gene	Length	Nh	*S*	θ	$\mu_{Relative}$	Tajima's *D*	Non-synonymous	Synonymous	$dN/dS_{Whitefish}$	$dN/dS_{Salmonids}$
ATP6	681	12	17	0.00504	1.11	−1.73	4	14	0.139	0.010
ATP8	165	2	1	0.00116	0.25	−1.02	0	1	NA	0.021
CYTB	1140	21	40	0.00675	1.48	−1.20	6	34	0.066	0.012
COX1	1548	14	24	0.00297	0.65	−0.90	0	24	0.000	0.008
COX2	690	9	11	0.00305	0.67	−1.20	1	10	0.028	0.019
COX3	783	7	13	0.00318	0.70	−0.48	1	12	0.013	0.020
ND1	972	13	36	0.00752	1.65	−0.64	3	35	0.012	0.018
ND2	1047	19	41	0.00772	1.69	−0.89	15	28	0.352	0.052
ND3	348	6	11	0.00607	1.33	−0.19	2	9	0.081	0.035
ND4	1380	19	41	0.00570	1.25	−1.05	3	38	0.025	0.024
ND4L	294	8	10	0.00654	1.43	−1.34	2	8	0.109	0.011
ND5	1836	20	60	0.00628	1.38	−0.63	10	50	0.056	0.039
ND6	519	7	14	0.00518	1.14	0.43	1	13	0.013	0.036
12S rRNA	947	8	7	0.00141	0.31	−1.84*	–	–	–	–
16S rRNA	1687	8	14	0.00159	0.35	−1.32	–	–	–	–
tRNAs	1555	15	19	0.00234	0.51	−1.09	–	–	–	–
Control region	1077	19	34	0.00628	1.38	−1.00	–	–	–	–
Mitogenome	16738	53	393	0.00456	1.00	−1.03	–	–	–	–

* Denotes significant Tajima's *D*.

3.2. Direction of nonsynonymous change

Overall ND2 contained a high number amino acid changes compared to the other genes. As such ND2 accounted for 50% (5/10) of all nonsynonymous changes fixed between American and European whitefish and 55.4% (11/21) when considering all changes supported in different mitogenome haplotypes (Fig. 2). Five of these were found in the terminal part of the ND2 gene (positions 328–348).

3.3. Analysis of positive selection

The results from TreeSAAP indicated that several significant physicochemical changes had occurred. The ND2 and ND5 genes showed the most significant changes (Table 2). Most of these were based upon recently derived mutations, which were only found within single individual sequences, with ND2 being the only exception.

The results of the MEME test found no positions involved in episodic selection (Table 2). The FUBAR test showed one position involved in diversifying positive selection (codon 331; empirical Bayes factor of 16.6). This position was located within the ND2 gene (Table 2) and showed as the only position two independent nonsynonymous mutations (Fig. 2) with one only observed in one sequence. Reducing the number of sequences in all three analyses to one from each evolutionary lineage or ecotype yielded similar results, showing no positive selected sites (data not shown).

3.4. Analysis of selection at the ND2 gene

For the ND2 dataset the *Whitefish vs. other* model showed a significantly better fit compared to the *Single ratio* model ($P > 0.001$) with a higher estimated dN/dS within the whitefish compared to the other salmonids (0.1919 vs. 0.0472; Table 3). For the other dataset constitution the 12 other mitochondrial genes the two models were not significantly different thus supporting an overall uniform dN/dS of 0.0233 for all branches ($P > 0.5$) (for phylogenetic tree see Fig. S2, supplemental material). MEME and FUBAR test using the entire salmonid phylogeny did not reveal any codons under episodic or diversifying positive selection in ND2.

Eight out of the 14 codon positions that showed amino acid change within the whitefish overlapped with positions showing nonsynonymous changes within the salmonid alignment (Fig. 3). This was significantly different from random expectations ($\chi^2 = 11.40$, $P < 0.01$). Of the last six positions four were only one codon position adjacent to a replaced codon positions in the salmonid alignment whereas position 24 and 247 were further apart (three and seven codons, respectively).

The other alignments also showed a significant overlap of amino acid changing positions with Scombridae overlapping at 7 positions ($\chi^2 = 4.39$, $P < 0.01$), Percidae at 12 positions ($\chi^2 = 7.14$, $P < 0.01$) and Anguillidae at 11 ($\chi^2 = 8.86$, $P < 0.01$). Simulation tests showed similar results (Note S1, supplemental material).

In total the whitefish ND2 gene showed 21% sequence identity to NuoN. The amino acid residues that have been suggested to participate in proton translocation (E133, K217, and K247 in NuoN) were conserved (Fig. 3). Most of the replaced changes were located

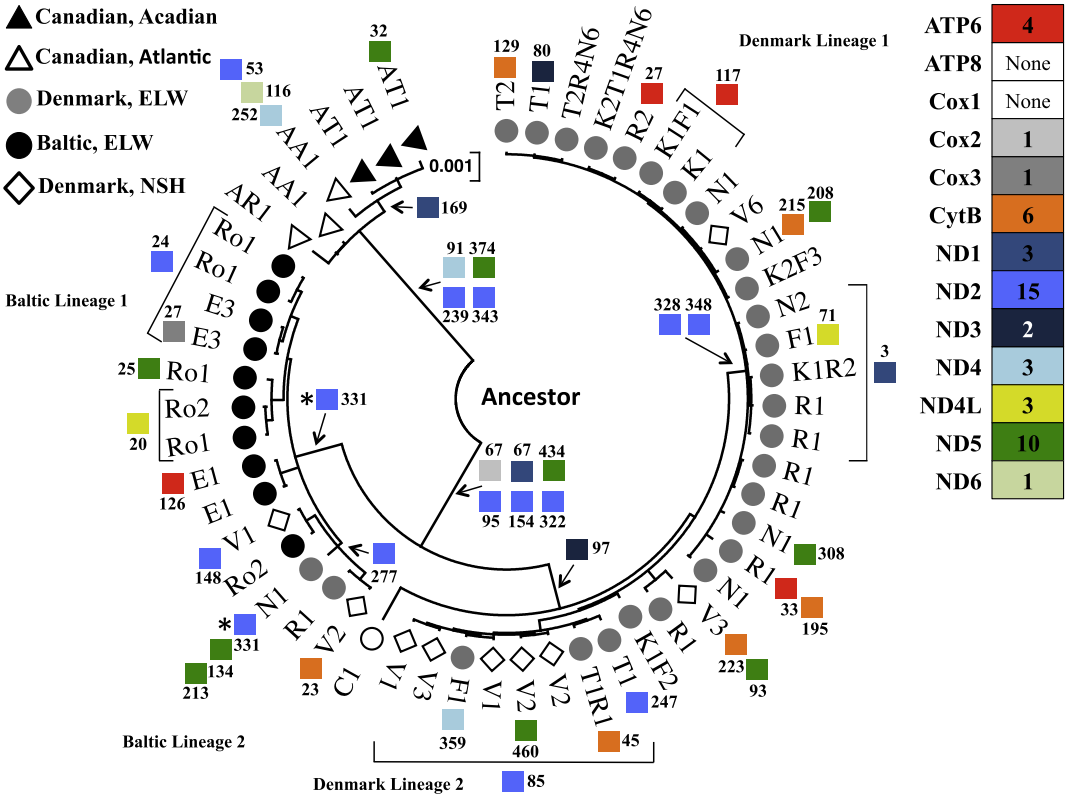

Fig. 2. Maximum-likelihood phylogeny of whitefish built using the whole mitogenome haplotype dataset. The different species and evolutionary lineages are denoted by different symbols at the tips. The direction of nonsynonymous change is visualized along the respective branches by colored squares with the adjacent numbers denoting the codon positions. The total number of nonsynonymous changes is shown within the legend on the right. '*' denotes the position under possible diversifying positive selection. Geographic localities are indicated as R = Ringkøbing fjord, N = Nissum fjord, V = Vidaa river, K = Kilen, F = Flynder lake, T = Tange lake, AW = Achterwasser, EC = Estonia coast, AT = Témiscouata lake, AA = Aylmer lake, AR = Ross lake. Values after the letters denote the number of identical haplotypes. For information on the amino acid changes see Table S2.

Table 2
Codons that candidates for being under positive selection, based on three selection tests.

Gene	Codon	TreeSAAP[a]			HyPhy		Minimum age (kyr)[b]
		AA-score[d]	*P*-value	Significant properties ($P < 0.001$)	FUBAR	MEME	
CYTB	223	7	<0.001	Power to be at the middle of alpha-helix	–	n.s.	0
ND2	53	8	<0.001	Equilibrium constant (ionization of COOH)	–	n.s.	–
	85	6	<0.001	Power to be at the C-terminal	–	n.s.	≅40
	328	8	<0.001	Equilibrium constant (ionization of COOH)	–	n.s	≅40
	331	8	<0.001	Equilibrium constant (ionization of COOH)	16.6	n.s.	≅132/≅16[c]
	348	6	<0.001	Power to be at the C-terminal	–	n.s.	≅40
ND3	80	6	<0.001	Power to be at the C-terminal	–	n.s.	–
ND4	252	6	<0.001	Power to be at the C-terminal	–	n.s.	–
ND5	32	7	<0.001	Isoelectric point	–	n.s.	–
	93	8	<0.001	Equilibrium constant (ionization of COOH)	–	n.s.	0
	134	8	<0.001	Equilibrium constant (ionization of COOH)	–	n.s.	–
	308	8	<0.001	Equilibrium constant (ionization of COOH)	–	n.s.	–
	460	8	<0.001	Equilibrium constant (ionization of COOH)	–	n.s.	–

[a] Results of TreeSAAP analyses. Only strong candidates for directional selection are shown (overall $P < 0.001$ and amino acid property score 6–8).
[b] Minimum ages of amino acid changes estimated from Jacobsen et al. (2012). Ages of 0 denotes cases where the nonsynonymous mutation is supported only by multiple identical haplotypes and '–' denotes cases where the nonsynonymous mutation is only found in a single individual sequence.
[c] This position shows two independent amino acid changes over time.
[d] Amino acid property score from TreeSAAP.

Table 3
CODEML analyses of *dN*/*dS* in whitefish compared to other salmonids for ND2 and the other 12 mitochondrial genes. Log-likelihoods, *dN*/*dS* and degrees of freedom (d.f.) are shown for all models together with results of the likelihood ratio tests.

Dataset	Models	LnL	*dN*/*dS* estimates	d.f.	χ^2 (2ΔL)	*P*-value
ND2	*Single ratio*	−6680.34	0.0498	1		
	Whitefish vs. others	−6675.10	0.1919/0.0472	2	10.48	<0.001
12 other mitochondrial genes	*Single ratio*	−54700.90	0.0233	1		
	Whitefish vs. others	−54700.28	0.0156/0.0234	2	1.23	$0.5 < P < 0.75$

outside or in the peripheral part of the transmembrane areas except for codons 247, 328 and 331 that clustered in buried domains. None were adjacent to the predicted proton translocation channels (Fig. 3). As a consequence of the low sequence identity, sites 343 and 348 could not be included in the model. Due to the low sequence identity, combined with low resolution of the X-ray structure used as a template for the whitefish ND2 model, it was not possible to assess the impact of the observed mutations on the function of ND2. No interactions were observed between the nonsynonymous changes and the ND5 homologue (subunits B and L of 3RKO) in the model (a Discovery Studio 4.1 Visualizer file of the ND2 model can be found in Supplemental material).

4. Discussion

In this study we present a comprehensive comparative analysis of mitogenome evolution in whitefish. Overall the synonymous substitutions in the mitochondrial coding genes evolve in a near neutral manner whereas the pattern for nonsynonymous substitutions is mostly consistent with strong purifying selection. The only gene showing possible evidence for positive selection is ND2. ND2 harbors an exceptionally high percentage of the total amino acid changes and shows higher *dN*/*dS* compared to the other genes and higher overall *dN*/*dS* in whitefish compared to the other salmonids. The result could underlie positive selection at some codon positions leading to a higher *dN*/*dS* in ND2 compared to the other genes. However, the physicochemical changes are overall of small effect and most changes fall within areas of the gene that are likely under less conservative evolutionary constraints. Moreover, codon-based test for positive selection revealed little evidence for positive selection. Thus we conclude that ND2 most likely is under relaxed purifying selection (here defined by *dN*/*dS* below one but higher compared to other genes due to less conservative evolutionary constraints).

4.1. Evolution of the mitochondrial regions in whitefish

Although recombination does not normally occur within the mitochondrial genome (Ballard and Whitlock 2004), different regions evolve at different rates (e.g. Doiron et al., 2002; Marshall et al., 2009; Whitehead, 2009). This is also observed in whitefish were the NADH genes, CYTB and control region evolve the fastest and the COX genes and RNAs the slowest (Table 1). This is the same overall pattern as observed in other fishes like Atlantic cod (*Gadus morhua*) (Marshall et al., 2009), killifish (*Fundulus heteroclitus*) (Whitehead, 2009) and chars (*Salvelinus fontinalis* and *S. alpinus*) (Doiron et al., 2002) and may be associated with the position of the genes in the mitogenome. CYTB, ND6, ND5, ND4 are found immediately downstream from the origin of H-strand replication (OriH) and ND1 and ND2 immediately upstream the origin of L-strand replication (OriL) (Nedbal and Flynn, 1998; Tanaka and Ozawa, 1994). Thus, during replication these genes will stay single stranded for longer time compared to the other genes rendering them more likely to accumulate mutations in the highly mutagenic environment of the mitochondrion (Marshall et al., 2009; Nedbal and Flynn, 1998).

Selection is likely also of importance, especially for understanding the differences in nonsynonymous rates observed between the mitochondrial genes (Meiklejohn et al., 2007; Mindell and Thacker, 1996). These rates vary as a likely consequence of differences in the strength of purifying selection due to functional constraints (e.g. Doiron et al., 2002; Meiklejohn et al., 2007). This is also evident from this study; although the numbers of neutrally evolving synonymous mutations are highly correlated with the length of the respective genes (Fig. 1(B)), reflecting overall neutral evolution, differences in the relative rates of nonsynonymous mutations exist between genes (Table 1, Fig. 1(C)). This is observed for both whitefish and other salmonids (Table 2) that overall show the same pattern, consistent with gene-specific differences in purifying selection.

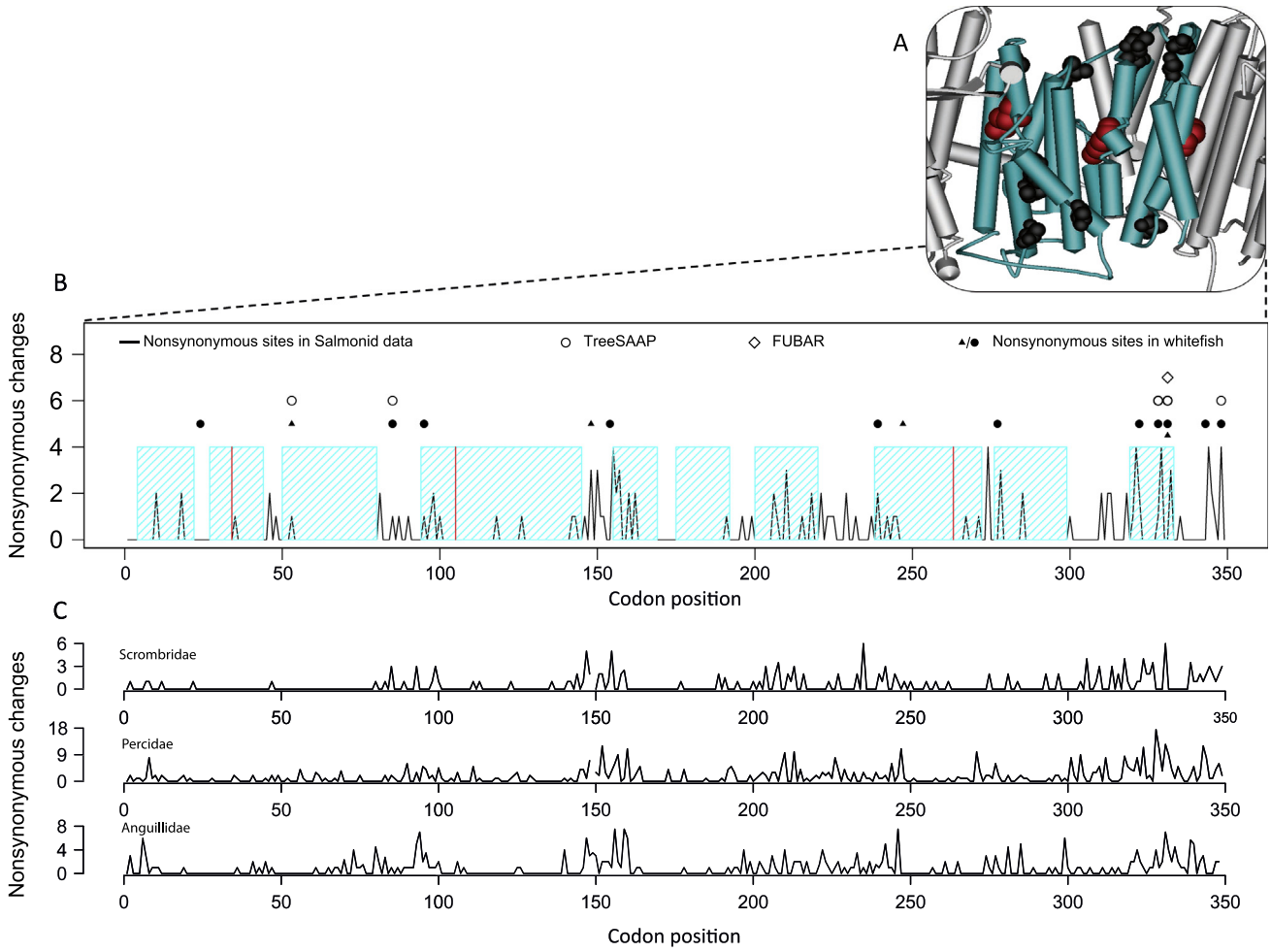

Fig. 3. Amino acid changes within the ND2 gene in whitefish. (A) The protein structure. ND2 is presented in blue, superimposed on chain D of PDB structure 3RKO. Gray structures correspond to the adjacent bacterial chains. Amino acids in red are the predicted proton translocation residues. Amino acids in black correspond to nonsynonymous sites observed in ⩾2 sequences. (B) Amino acid change along the gene. The black line shows the total number of estimated nonsynonymous changes within the salmonid alignment while the black symbols represent the positions of amino acid changes in whitefish. Dots denote amino acid changes supported by >2 sequences while the triangles denote unshared changes only found in a single sequence. Codon positions found to be under positive selection are shown by different symbols and correspond to the tests presented in the top of the figure. The blue squares denote the transmembrane domains predicted from the protein modeling and the red line the amino acid residues that have been suggested to participate in proton translocation. (C) Number of estimated amino acid changes within the ND2 gene in Scombridae, Percidae and Anguillidae.

4.2. Investigation of positive selection in the mitochondrial genes in whitefish

Given the phenotypic and physiological differences relating to energetics observed among several of the whitefish species and lineages analyzed here, positive selection within the mitochondrial genes seems likely. This finds further support as several studies have reported possible adaptive selection at the regulatory level in genes involved in the OXPHOS pathway (Derome et al., 2006; Evans and Bernatchez, 2012; Jensen et al., 2013). For example, ND5, CYTB and ATP8 show significant upregulation in dwarf compared to normal ALW (Evans and Bernatchez, 2012). Also between the Danish NSH and ELW differences in gene expression has been observed in the nuclear ATP synthase subunit g (Jensen et al., 2013), which together with e.g. the mitochondrial ATP6 and ATP8 genes constitutes ATP synthase (complex V) of the OXPHOS pathway (Saraste, 1999). Another gene involved in ATP synthase, the ATP synthase subunit c shows significant upregulation in dwarf versus normal ALW forms (Derome et al., 2006). However, although these studies support the possibility for co-adaptations and positive selection occurring at the functional level within the mitochondrial coding genes in whitefish, little evidence for positive selection was found (Table 2). As for most studied species (e.g. da Fonseca et al., 2008; Sun et al., 2011), analyses of gene specific *dN*/*dS* all showed values <1 consistent with purifying selection (Table 1). Furthermore, tests for episodic and diversifying selection showed no codons under positive selection; the only exception was ND2, which is discussed below. TreeSAAP analyses for overall significant physicochemical changes showed some significant amino acid changes. However, most of them concerned changes in the terminal branches (Table 2, Fig. 2) and were often only found within single individual sequences. Thus although these changes may be adaptive, the pattern is more consistent with relaxed purifying selection, as supported by slightly elevated *dN*/*dS* in the terminal branches (Jacobsen et al., 2012).

In total the lack of clear evidence for positive selection at the coding level is in contrast to the studies reporting adaptive selection at the regulatory level of mitochondrial or nuclear interactor genes (Derome et al., 2006; Evans and Bernatchez, 2012; Jensen et al., 2013). This fact indicates an important role of regulatory differences in explaining the ecological and biological differences among the newly evolved whitefish species and ecotypes examined here.

4.3. Distinguishing between relaxed purifying selection and positive selection in ND2 whitefish

The NADH dehydrogenase complex is the first and largest of the five enzyme complexes that constitute the oxidative phosphorylation pathway (Ballard and Whitlock, 2004). It receives electrons from the oxidation of NADH and provides electrons for reduction of quinone to quinol, which leads to the translocation of four protons across the inner membrane. This generates an electrochemical proton gradient that drives the later stages of energy production (Ballard and Whitlock, 2004; da Fonseca et al., 2008). ND2, ND4 and ND5 are suggested to be the actual proton pumping devices (Brandt 2006) and amino acid changes here may therefore potentially be adaptive. Positive selection at ND2 has been suggested in several organisms by e.g. conducting TreeSAAP analyses of radical amino acid property changes in Atlantic herring (*Clupea harengus*) and Pacific salmon (*Oncorhynchus* sp.) (Garvin et al., 2011; Teacher et al., 2012) and branch-site likelihood ratio tests for significantly elevated codon-specific *dN*/*dS* in Chinese snub-nosed monkeys (*Rhinopithecus* sp.) (Yu et al., 2011). Moreover outlier testing and logistic regression between metal contamination level and alleles showed a link between nonsynonymous change in ND2 and cobber contamination in yellow perch (*Perca flavescens*) (Belanger-Deschenes et al., 2013).

In whitefish, ND2 shows several patterns compatible with positive selection. A much higher *dN*/*dS* ratio is observed compared to the other genes and compared to other salmonids (Tables 1 and 3, Figs. 1(C) and 3(A) and (B)) and over 50% of all amino acid differences found in more than one haplotype in whitefish are found in ND2 (Fig. 2). Moreover, as the only gene, TreeSAAP analysis demonstrates significant changes of four replacement mutations of older (non-terminal) nature. Finally, ND2 shows one codon position under possible diversifying positive selection. However, although these results could reflect positive selection it may not be the most parsimonious explanation. If positive selection had occurred within specific lineages we would expect them to show elevated nonsynonymous rates (Murrell et al., 2012). Although PAML analysis showed higher *dN*/*dS* within the whitefish there is no evidence for episodic selection acting on the single codon positions (Table 2). Moreover, in codon 331 showing potential diversifying positive selection, one of the two amino acid changes is only observed in one sequence (Fig. 2), which could potential relate to relaxed purifying selection (Jacobsen et al., 2012). More importantly, most of the amino acid changes in whitefish are situated outside the transmembrane areas and none include the residues suggested to participate in proton translocation (Fig. 3). In a mammal study, TreeSAAP analyses showed that ND2 loop areas in general harbor more radical amino acid changes than transmembrane domains, which suggests lower functional constraints of these areas (da Fonseca et al., 2008). A similar finding is also reported for killifish (Whitehead, 2009) and is consistent with the putative proton-pumping function of the transmembrane domains (Brandt, 2006; Efremov and Sazanov, 2011). It is therefore likely that these regions are under relaxed functional constraints. This is further supported by the significant overlap observed between regions of nonsynonymous change in whitefish and other salmonids, as well as the three other groups of fish investigated (Fig. 3). These changes are in most cases located outside transmembrane areas. The only major exception is within the last predicted transmembrane domain. However, this part of ND2 shows extensive nonsynonymous change in all investigated fishes, which again suggests relaxed constraints in this region.

Although the patterns in whitefish appear to be best explained by relaxed purifying selection, one question remains. Why is the nonsynonymous substitution rate so much higher for the ND2 gene compared to the other mitochondrial coding genes in whitefish? We propose four possibilities. (1) ND2 simply has fewer constraints compared to the other mitochondrial genes. This finds some support, as several studies report ND2 to contain some of the highest proportions of radical physicochemical amino acid changes (da Fonseca et al., 2008; Teacher et al., 2012; Whitehead, 2009) and generally shows high rates of nonsynonymous changes (Doiron et al., 2002). A higher nonsynonymous rate of ND2 is supported in this study as PAML analyses showed higher *dN*/*dS* for the ND2 gene compared to the other genes (Tables 1 and 3). However, not all studies identify ND2 as a distinct outlier and other mechanisms may be acting. (2) It is possible that the timescale investigated in this study leads to stochasticity in rates compared to studies conducted over longer timespans of evolution. The estimated time of the most recent common ancestor (TMRCA) of all whitefish sequences examined here, is between 1.33 (95%HPD: 1.14–1.50) and 0.43 Myr (95%HPD: 0.37–0.49), depending on the substitution rate used (Jacobsen et al., 2012). This could explain the lower *dN*/*dS* observed for the other salmonids (Table 2, Fig. 3) that covers much deeper evolutionary histories with a TMRCA of 59.1 Myr (CI: 58.1–63.2 Myr) (Crete-Lafreniere et al., 2012). Here purifying selection may have had more time to act against slightly deleterious mutation fixed by genetic drift, which would lead to a lower *dN*/*dS*. However, given *dN*/*dS* estimated for the 12 other genes in whitefish is not significantly higher compared to the other salmonids, purifying selection seems inadequate for explaining the higher *dN*/*dS* within ND2 in whitefish. Another possibility (3) is that drift combined with selection has shaped the pattern observed at ND2, according to the compensatory-draft model presented by Oliveira et al. (2008). The model states that mild-deleterious replacement mutations fixed by e.g. genetic drift will lead to positive selection for compensatory mutations. Given the non-recombining nature of the mitogenome this will lead to genetic sweeps and thus possible fixation of additional deleterious mutations, which again will allow for additional sweeps of compensatory changes. Over time this may result in high relative rates of nonsynonymous substitutions in affected genes, as proposed for ATP6 and ATP8 in parasitic wasps (*Nasonia* sp.) (Oliveira et al., 2008). However, if such a mechanism were at work in whitefish ND2, we would expect to find evidence of episodic selection at closely situated codons. However, MEME tests using the whitefish and salmonids datasets did not reveal any sites under episodic selection. Nonetheless, we observed a high number of nonsynonymous substitutions in whitefish in the terminal part of the ND2 gene, which harbors 5 of the 11 nonsynonymous substitutions found in $\geqslant 2$ sequences (Fig. 2). Four of these occur along the branches within the European whitefish clade and thus match expectations according to the compensatory-draft model. However, this region shows high numbers of nonsynonymous changes in all investigated taxa and the nonsynonymous positions observed in whitefish, overlaps with positions in salmonids. As such this favors relaxed purifying selection. A final possibility (4) is that ND2 serves other functions outside the oxidative phosphorylation chain. This has been demonstrated by Gingrich et al. (2004) whom showed ND2 to be present in the brain. Specifically, ND2 served as an adaptor anchoring Src to the N-methyl-D-aspartate receptor (NMDAR) in the brain synapses. Regulation of NDMAR by Src-ND2 is thought to be essential to long-term potential (LTP) (Gingrich et al., 2004), the main mechanism explaining learning and memory in animals (Kandel, 2001). The anchoring domain belongs to the 239–321 codon in human ND2, a region that also shows change in whitefish (position 239, 247 and 277, Fig. 2, Table S2). In a recent study, Belanger-Deschenes et al. (2013) demonstrated that cobber-contaminated Canadian populations of Yellow Perch (*Perca flavescens*) show amino acid change at codon position 278 in ND2. This position shows a radical change in amino acid property from threonine (slightly hydrophilic and polar) to alanine (hydrophilic and non-polar). As cobber is known to block

NMDAR, which may lead to inhibition of LTP, they suggested that the amino acid change could be adaptive in cobber-contaminated populations. However, this additional function of ND2 cannot explain the higher rate of ND2 in whitefish. The adaptor region only covers three of the 15 nonsynonymous sites in ND2 (Fig. 2). Of these, position 247 is only found within one sequence and position 277 is the only one located well within the region. Moreover, the haplotypes with this mutation are not associated to known differences in habitat, contamination level or biology that could otherwise be associated with learning. In fact, the mutation is found in one evolutionary lineage representing four different populations inhabiting either brackish (Achterwasser in the German part of the Baltic Sea and Nissum and Ringkøbing fjord on the west coast of Denmark) or marine environments (Vidaa river) with none showing a record of severe pollution. Furthermore, in all populations the mutation is only found in low frequency (ranging from 12.5% [1/8] to 5% [1/20] in the Rostock population and the Ringkøbing and Nissum fjord population respectively, data not shown), which favors genetic drift over positive selection. Lastly, ND5 interacts extensively with ND2 (e.g. Garvin et al., 2011), which could promote selection within ND2. However, in our model we found no interactions between the nonsynonymous sites in ND2 and the ND5 homologue (subunits B and L of 3RKO in the model; Supplemental material). Although this could underlie uncertainties in the model we find no evidence for positive selection and the possibility for co-adaptation between the two subunits seems unlikely.

Given the potential explanations discussed above we find that the high non-synonymous rate observed in ND2 in whitefish is most parsimonious with relaxed purifying selection, relating to regions under less conservative evolutionary constraints, combined with an overall high mutation rate of ND2. However, given possible unknown functions of ND2, positive selection cannot be ruled out entirely.

5. Conclusion

We investigated the evolutionary patterns and especially the possibility for positive selection in mitochondrial genes of 107 whitefish. Although the dataset compromised several unique species and evolutionary lineages that show distinct phenotypic differences relating to metabolism we found no strong support for positive selection at the coding level. The most likely candidate for positive selection was the ND2 gene that showed highly elevated *dN*/*dS* compared to the other genes. Although *dN*/*dS* was below 1, in favor of purifying selection, the result could underlie positive selection acting at the few codon positions leading to a higher *dN*/*dS* compared to the other genes. However, evidence for codon-specific positive selection in ND2 was weak. Furthermore, the positions of amino acid change in ND2 in whitefish overlapped significantly with nonsynonymous positions in other salmonids, as well as three other groups of fish (Scombridae, Percidae and Anguillidae) and were not associated with changes in three key positions responsible for proton translocation. The changes were generally found within the loop regions, which have been suggested to be under relaxed constraints in mammals and thus the pattern found in ND2 seems best explained by relaxed purifying selection.

In total, our study emphasizes the challenges when distinguishing between positive and relaxed selection and highlights the importance of using different analyses when assessing selection at the mitogenome level.

Acknowledgments

The authors acknowledge the Danish Nature Agency – Denmark (Naturstyrelsen), the EU-LIFE Programme (LIFE05 NAT/DK/000153 – Actions for Houting) and the Danish Council for Independent Research Natural Sciences – Denmark (grant no. 1323-00158 to MMH) for funding. We moreover thank Marcus Thomas Pius Gilbert for comments and Virginia Settepani for helping with the PAML analyses.

Appendix. A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.11.008.

References

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744.
Bazin, E., Glemin, S., Galtier, N., 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312, 570–572.
Belanger-Deschenes, S., Couture, P., Campbell, P.G.C., Bernatchez, L., 2013. Evolutionary change driven by metal exposure as revealed by coding SNP genome scan in wild yellow perch (*Perca flavescens*). Ecotoxicology 22, 938–957.
Bernatchez, L., Dodson, J.J., 1990. Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA restriction analysis. Evolution 44, 1263–1271.
Bernatchez, L., Dodson, J.J., 1994. Phylogenetic-relationships among palearctic and nearctic whitefish (*Coregonus* sp) populations as revealed by mitochondrial-DNA variation. Can. J. Fish. Aquat. Sci. 51, 240–251.
BioMatters, 2012. Geneious version 5.4.6. Available at <http://www.geneious.com>.
Birky, C.W., Maruyama, T., Fuerst, P., 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics 103, 513–527.
Brandt, U., 2006. Energy converting NADH: quinone oxidoreductase (Complex I). Annu. Rev. Biochem. 75, 69–92.
Campbell, M.A., Lopez, J.A., Sado, T., Miya, M., 2013. Pike and salmon as sister taxa: detailed intraclade resolution and divergence time estimation of Esociformes plus Salmoniformes based on whole mitochondrial genome sequences. Gene 530, 57–65.
Crete-Lafreniere, A., Weir, L.K., Bernatchez, L., 2012. Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. PLoS ONE 7, e46662.
da Fonseca, R.R., Johnson, W.E., O'Brien, S.J., Ramos, M.J., Antunes, A., 2008. The adaptive evolution of the mammalian mitochondrial genome. BMC Genomics 9.
Dalziel, A.C., Moyes, C.D., Fredriksson, E., Lougheed, S.C., 2006. Molecular evolution of cytochrome c oxidase in high-performance fish (Teleostei: Scombroidei). J. Mol. Evol. 62, 319–331.
Derome, N., Duchesne, P., Bernatchez, L., 2006. Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchill) ecotypes. Mol. Ecol. 15, 1239–1249.
Doiron, S., Bernatchez, L., Blier, P.U., 2002. A comparative mitogenomic analysis of the potential adaptive value of arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill). Mol. Biol. Evol. 19, 1902–1909.
Efremov, R.G., Sazanov, L.A., 2011. Structure of the membrane domain of respiratory complex I. Nature 476, 414–420.
Evans, M.L., Bernatchez, L., 2012. Oxidative phosphorylation gene transcription in whitefish species pairs reveals patterns of parallel and nonparallel physiological divergence. J. Evol. Biol. 25, 1823–1834.
Fontanillas, P., Depraz, A., Giorgi, M.S., Perrin, N., 2005. Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, *Crocidura russula*. Mol. Ecol. 14, 661–670.
Foote, A.D., Morin, P.A., Durban, J.W., Pitman, R.L., Wade, P., Willerslev, E., Gilbert, M. T.P., da Fonseca, R.R., 2011. Positive selection on the killer whale mitogenome. Biol. Lett. 7, 116–118.
Gagnaire, P.A., Normandeau, E., Bernatchez, L., 2012. Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American eels. Mol. Biol. Evol. 29, 2909–2919.
Gagnaire, P.A., Pavey, S.A., Normandeau, E., Bernatchez, L., 2013. The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by rad sequencing. Evolution 67, 2483–2497.
Galtier, N., Nabholz, B., Glemin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541–4550.
Garvin, M.R., Bielawski, J.P., Gharrett, A.J., 2011. Positive Darwinian selection in the piston that powers proton pumps in complex I of the mitochondria of pacific salmon. PLoS ONE 6, e24127.
Gingrich, J.R., Pelkey, K.A., Fam, S.R., Huang, Y.Q., Petralia, R.S., Wenthold, R.J., Salter, M.W., 2004. Unique domain anchoring of Src to synaptic NMDA receptors via the mitochondrial protein NADH dehydrogenase subunit 2. Proc. Natl. Acad. Sci. U.S.A. 101, 6237–6242.
Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. PAST: paleontological statistics software package for education and data analysis. Palaeontol. Electron. 4, 9, <http://palaeo-electronica.org/2001_2001/past/issue2001_2001.htm>.
Hansen, M.M., Fraser, D.J., Als, T.D., Mensberg, K.L.D., 2008. Reproductive isolation, evolutionary distinctiveness and setting conservation priorities: the case of

European lake whitefish and the endangered North Sea houting (Coregonus spp.). BMC Evol. Biol. 8.
Hansen, M.M., Nielsen, E.E., Mensberg, K.L.D., 2006. Underwater but not out of sight: genetic monitoring of effective population size in the endangered North Sea houting (*Coregonus oxyrhynchus*). Can. J. Fish. Aquat. Sci. 63, 780–787.
Himberg, M.K.-J., Lehtonen, H., 1995. Systematics and nomenclature of coregonid fishes, particularly in Northwest Europe. Adv. Limnol. 46, 39–47.
Jacobsen, M.W., Hansen, M.M., Orlando, L., Bekkevold, D., Bernatchez, L., Willerslev, E., Gilbert, M.T.P., 2012. Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and ecologically distinct European whitefish (*Coregonus* spp.). Mol. Ecol. 21, 2727–2742.
Jacobsen, M.W., Pujolar, J.M., Gilbert, M.T.P., Moreno-Mayar, J.V., Bernatchez, L., Als, T.D., Lobon-Cervia, J., Hansen, M.M., 2014. Speciation and demographic history of Atlantic eels (*Anguilla anguilla* and *A. rostrata*) revealed by mitogenome sequencing. Heredity 113, 432–442.
Jacobsen, M.W., Pujolar, J.M., Hansen, M.M., 2015. Relationship between amino acid changes in mitochondrial ATP6 and life history variation in Anguillid eels. Biol. Lett. 11, 20150014.
Jensen, A., Nielsen, H., Ejbye-Ernst, M., 2003. National forvaltningsplan for snæbel. In: Report from Miljøministeriet, Skov og Natarstyrrelsen, Sønderlands Amt og Ribe Amt.
Jensen, L.F., Poulsen, S.B., Svendsen, J.C., 2013. Fish and chips – undersøgelser af snæblens (*Coregonus oxyrinchus*) tilpasninger til det lokale miljø. Sjæklen, 104–115.
Jonckheere, A.I., Smeitink, J.A.M., Rodenburg, R.J.T., 2012. Mitochondrial ATP synthase: architecture, function and pathology. J. Inherit. Metab. Dis. 35, 211–225.
Kandel, E.R., 2001. Neuroscience – the molecular biology of memory storage: a dialogue between genes and synapses. Science 294, 1030–1038.
Kottelat, M., Freyhof, J., 2007. Handbook of European Freshwater Fishes, Berlin, Germany.
Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.
Lu, G., Basley, D.J., Bernatchez, L., 2001. Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. Mol. Ecol. 10, 965–985.
Marshall, H.D., Coulson, M.W., Carr, S.M., 2009. Near neutrality, rate heterogeneity, and linkage govern mitochondrial genome evolution in Atlantic cod (*Gadus morhua*) and other gadine fish. Mol. Biol. Evol. 26, 579–589.
Meiklejohn, C.D., Montooth, K.L., Rand, D.M., 2007. Positive and negative selection on the mitochondrial genome. Trends Genet. 23, 259–263.
Mindell, D.P., Thacker, C.E., 1996. Rates of molecular evolution: phylogenetic issues and applications. Annu. Rev. Ecol. Syst. 27, 279–303.
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.
Murrell, B., Moola, S., Mabona, A., Mabona, A., Weighill, T., Sheward, D., Pond, S.L.K., Scheffler, K., 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol. Biol. Evol. 30, 1196–1205.
Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Pond, S.L.K., 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 8, e1002764.
Nedbal, M.A., Flynn, J.J., 1998. Do the combined effects of the asymmetric process of replication and DNA damage from oxygen radicals produce a mutation-rate signature in the mitochondrial genome? Mol. Biol. Evol. 15, 219–223.
Oliveira, D.C.S.G., Raychoudhury, R., Lavrov, D.V., Werren, J.H., 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp Nasonia (Hymenoptera: Pteromalidae). Mol. Biol. Evol. 25, 2167–2180.
Østbye, K., Amundsen, P.A., Bernatchez, L., Klemetsen, A., Knudsen, R., Kristoffersen, R., Naesje, T.F., Hindar, K., 2006. Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times. Mol. Ecol. 15, 3983–4001.
Østbye, K., Bernatchez, L., Naesje, T.F., Himberg, K.J.M., Hindar, K., 2005. Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. Mol. Ecol. 14, 4371–4387.
Piganeau, G., Eyre-Walker, A., 2009. Evidence for variation in the effective population size of animal mitochondrial DNA. PLoS ONE 4, e4396.
Pigeon, D., Chouinard, A., Bernatchez, L., 1997. Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). Evolution 51, 196–205.
Pond, S.L.K., Frost, S.D.W., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22, 1208–1222.
Rand, D.M., Haney, R.A., Fry, A.J., 2004. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol. Evol. 19, 645–653.
Rogers, S.M., Bernatchez, L., 2005. Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). Mol. Ecol. 14, 351–361.
Rogers, S.M., Bernatchez, L., 2007. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp Salmonidae) species pairs. Mol. Biol. Evol. 24, 1423–1438.
Saraste, M., 1999. Oxidative phosphorylation at the fin de siecle. Science 283, 1488–1493.
Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381–3385.
Smeitink, J., van den Heuvel, L., DiMauro, S., 2001. The genetics and pathology of oxidative phosphorylation. Nat. Rev. Genet. 2, 342–352.
Sun, Y.B., Shen, Y.Y., Irwin, D.M., Zhang, Y.P., 2011. Evaluating the roles of energetic functional constraints on teleost mitochondrial-encoded protein evolution. Mol. Biol. Evol. 28, 39–44.
Tajima, F., 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.
Tajima, F., 1996. The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics 143, 1457–1465.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Tanaka, M., Ozawa, T., 1994. Strand asymmetry in human mitochondrial DNA mutations. Genomics 22, 327–335.
Teacher, A.G., Andre, C., Merila, J., Wheat, C.W., 2012. Whole mitochondrial genome scan for population structure and selection in the Atlantic herring. BMC Evol. Biol. 12.
Trudel, M., Tremblay, A., Schetagne, R., Rasmussen, J.B., 2001. Why are dwarf fish so small? An energetic analysis of polymorphism in lake whitefish (*Coregonus clupeaformis*). Can. J. Fish. Aquat. Sci. 58, 394–405.
Whitehead, A., 2009. Comparative mitochondrial genomics within and among species of killifish. BMC Evol. Biol. 9.
Woolley, S., Johnson, J., Smith, M.J., Crandall, K.A., McClellan, D.A., 2003. TreeSAAP: selection on amino acid properties using phylogenetic trees. Bioinformatics 19, 671–672.
Yang, Z.H., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.
Yasuike, M., Jantzen, S., Cooper, G.A., Leder, E., Davidson, W.S., Koop, B.F., 2010. Grayling (Thymallinae) phylogeny within salmonids: complete mitochondrial DNA sequences of *Thymallus arcticus* and *Thymallus thymallus*. J. Fish Biol. 76, 395–400.
Yu, L., Wang, X.P., Ting, N., Zhang, Y.P., 2011. Mitogenomic analysis of Chinese snub-nosed monkeys: evidence of positive selection in NADH dehydrogenase genes in high-altitude adaptation. Mitochondrion 11, 497–503.