Preliminary assessment of adaptive evolution of mitochondrial protein coding genes in darters (Percidae: Etheostomatinae)

Leos G. Kral, Sara Watson

Department of Biology, University of West Georgia, Carrollton, GA, 30118, USA

Abstract

**Background:** Mitochondrial DNA of vertebrates contains genes for 13 proteins involved in oxidative phosphorylation. Some of these genes have been shown to undergo adaptive evolution in a variety of species. This study examines all mitochondrial protein coding genes in 11 darter species to determine if any of these genes show evidence of positive selection.

**Methods:** The mitogenome from four darter was sequenced and annotated. Mitogenome sequences for another seven species were obtained from GenBank. Alignments of each of the protein coding genes were subject to codon-based identification of positive selection by Selecton, MEME and FEL.

**Results:** Evidence of positive selection was obtained for six of the genes by at least one of the methods. *CYTB* was identified as having evolved under positive selection by all three methods at the same codon location.

**Conclusions:** Given the evidence for positive selection of mitochondrial protein coding genes in darters, a more extensive analysis of mitochondrial gene evolution in all the extant darter species is warranted.

**Keywords**

Mitogenome, mtDNA, adaptive evolution, positive selection, Etheostoma, Percina

**Corresponding author:** Leos G. Kral (lkral@westga.edu)

**Author roles:** Kral LG: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation; Watson S: Data Curation, Formal Analysis, Investigation, Methodology

**Competing interests:** No competing interests were disclosed.

**Grant information:** This study was supported by a Faculty Research Grant funded by the University of West Georgia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Copyright:** © 2019 Kral LG and Watson S. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Kral LG and Watson S. Preliminary assessment of adaptive evolution of mitochondrial protein coding genes in darters (Percidae: Etheostomatinae) [version 1; peer review: 1 not approved] F1000Research 2019, 8:464 (https://doi.org/10.12688/f1000research.17552.1)

First published: 14 Apr 2019, 8:464 (https://doi.org/10.12688/f1000research.17552.1)
Introduction
Mitochondria provide cells with nearly all energy as a result of oxidative phosphorylation (OXPHOS). Vertebrate mitochondria are primarily maternally inherited and contain mitochondrial DNA (mtDNA), which contains 13 protein coding genes, the products of which contribute, along with nuclear-encoded proteins (Sunnucks et al., 2017), to the formation of the mitochondrial OXPHOS machinery (Ingman & Gyllensten, 2006; Ladoukakis & Zouros, 2017). While purifying selection acts on deleterious mtDNA mutations (Burr et al., 2018), James et al. (2016) utilized a variety of McDonald-Kreitman type analyses of a large number of animal species to show that mtDNA evolution is dominated by slightly deleterious mutations but also undergoes a significant amount of adaptive evolution. Several studies of the evolution of mtDNA-encoded mitochondrial protein coding genes both within fish species (Consuegra et al., 2015; Teacher et al., 2012) and between fish species (D’Anatro et al., 2017; Garvin et al., 2011; Zhang & Broughton, 2015) identified positive selection acting on some of these genes. This study utilized all currently available mitochondrial genome sequences of darter species to determine if positive selection can be detected in any of the mitochondrial protein coding genes. Such positive selection would imply that adaptation of the OXPHOS machinery to differing environmental niches had occurred in at least some lineages during darter speciation.

Methods
Mitochondrial genomes
All mitochondrial genome sequences utilized in this study were obtained from GenBank except those from Etheostoma chuckwachatte, Etheostoma jessiae, Etheostoma spectabile, Etheostoma tallapoosae and Percina crypta. The E. spectabile sequence was obtained from Rachel Moran, University of Illinois. Whole genome sequencing was performed on E. chuckwachatte, E. jessiae, E. tallapoosae and P. crypta genomic DNA (purified from fin/muscle tissue with the Qiagen DNeasy Blood and Tissue Kit) by the Georgia Genomic Facility at the University of Georgia utilizing the Illumina NexSeq (PE 150) or MiSeq (PE 250) sequencer. A subset of sequence reads from each species, sufficient for at least 30-fold coverage, was aligned to one of the GenBank obtained darter mitochondrial genomes utilizing the Map to Reference function in Geneious 9.1.8 software (Biomatters Ltd., Auckland, New Zealand). Annotated darter mitochondrial genome sequences from GenBank were aligned with the newly assembled darter mitochondrial genome sequences and annotations were transferred and manually adjusted where necessary with Geneious software.

Phylogenetic analysis
Phylogenetic analysis was carried out on all 13 concatenated and aligned mitochondrial protein coding sequences of the 11 darter species as well as of Perca flavescens, Sander vitreus and Anarhichas minor (outgroup). Bayesian inference (BI) analysis was performed with MrBayes (Ronquist & Huelsenbeck, 2003), as implemented in Geneious software but utilizing optimal partitioning and optimal substitution models determined by PartitionFinder v2.1.1 (AICc, greedy algorithm, MrBayes specific models) (Guindon et al., 2010; Lanfear et al., 2012; Lanfear et al., 2017). The rooted darter sub-tree (Figure 1) was extracted with MEGA 7.0.26 software (Kumar et al., 2016). Aligned darter mitochondrial protein coding genes were subject to site-specific tests for positive selection utilizing the Selecton implementation of an empirical Bayes approach (Doron-Faigenboim et al., 2005; Stern et al., 2007) and also the MEME and FEL methods implemented on the Datamonkey webserver (Delport et al., 2010). The darter tree was used in all three of these analyses.

Results and discussion
Identification of codons under positive selection
Codons under positive selection were identified in seven of the thirteen mitochondrial protein coding genes by at least one of the methods utilized. Specifically, as shown in Table 1, the MEME method identified two codons in COX1, two codons in CYTB, three codons in ND2, and one codon each in ND3 and in ND5. The FEL method identified one codon in COX1, CYTB and ND3. The codons identified as being under positive selection by the FEL method were also identified as being under positive selection by MEME method. Specifically, codon 489 in COX1, codon 96 in CYTB and codon 9 in ND3. As shown in Table 2, the Selecton method identified one codon in ATP6, one codon in COX3, one codon in CYTB and two codons in ND5. While codons under positive selection were identified in ND5 by MEME and Selecton, these methods did not identify the same codons. Specifically, codon 32 was identified in ND5 by MEME and codons 479 and 573 were identified by Selecton. Only codon 96 in CYTB was identified as being under positive selection by all three methods.

Analysis of sites under positive selection
While the Bayesian estimates of the Ka/Ks ratio of the sites identified by Selecton are greater than 1, the lower bounds of the confidence intervals defined by the 5th and 95th percentiles of the posterior distribution inferred for these sites are less than 1. Therefore, the reliability of positive selection of these sites identified by Selecton is not very strong. However, positive selection of the four proteins identified by Selecton has been found to be significant (p = 0.001) by the likelihood ratio test of log-likelihood for model M8 (allowing positive selection) versus model M8a (not allowing positive selection).

Drivers of selection
These results indicate that it is likely that the evolution of various darter lineages included positive selection of at least some of the mitochondrially encoded OXPHOS machinery variants. Presumably, this positive selection would have been driven by adaptation to specific environmental factors. As an example of mitochondrial adaptation to environmental factors, Ma et al. (2015) found that adaptation of Chinese glyptosternoid fishes to high-elevation of the Tibetan Plateau is correlated to signals of positive selection of the Cox1 gene. That adaptation to hypoxia is correlated to physiological adaptation of the OXPHOS machinery has been demonstrated by O2 binding kinetics of cytochrome c oxidase (COX) in hypoxia tolerant vs. intolerant
Figure 1. Darter phylogenetic tree produced using Bayesian estimation (Mr. Bayes) of concatenated mitochondrial protein coding genes. Branch lengths measured in number of substitutions per site. Posterior probabilities of all branches are greater than 0.91. Pm, Percina macrolepida; Pc, Percina crypta; Ecw, Etheostoma chuckwachatte; Ec, Etheostoma caeruleum; Er, Etheostoma radiosum; Eo, Etheostoma okaloosae; Es, Etheostoma spectabile; Ejs, Etheostoma jessiae; En, Etheostoma nigrum; Ez, Etheostoma zonale; Et, Etheostoma tallapoosae.

Table 1. Genes and sites inferred to be under positive selection by MEME and FEL. Default significance cutoff value of p<0.1 was used.

<table>
<thead>
<tr>
<th>Gene</th>
<th>MEME</th>
<th>FEL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Site</td>
<td>p value</td>
</tr>
<tr>
<td>COX1</td>
<td>3</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>489</td>
<td>0.06</td>
</tr>
<tr>
<td>CYTB</td>
<td>82</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.02</td>
</tr>
<tr>
<td>ND2</td>
<td>119</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>327</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>339</td>
<td>0.05</td>
</tr>
<tr>
<td>ND3</td>
<td>9</td>
<td>0.06</td>
</tr>
<tr>
<td>ND5</td>
<td>32</td>
<td>0.08</td>
</tr>
</tbody>
</table>
sculpin species (Lau et al., 2017). In this study, several non-synonymous substitutions in COX3 were identified by in silico analysis as candidates for explaining the adaptive variation of COX O$_2$ binding. That mitochondrial haplotype variants are subject selection by environmental factors has been demonstrated experimentally in Drosophila. When a mixed population of “cold adapted” haplogroups and “warm adapted” haplogroups were maintained for a number of generations at differing temperatures, the frequency of the “cold adapted” haplogroup increased at the cooler temperature and decreased at the warmer temperature (Lajbner et al., 2018).

Given the preliminary results indicating positive mitochondrial gene selection in darters, it would seem that a more expansive analysis of the evolution of mitochondrial protein coding genes in the approximately 200 extant darter species (Near et al., 2011) may be warranted to identify lineage specific adaptations and, potentially, their correlations to relevant life history traits.

**Data availability**

**Underlying data**

All mitogenome sequences are available in GenBank.


<table>
<thead>
<tr>
<th>Gene</th>
<th>Position</th>
<th>Ka/Ks</th>
<th>CI</th>
<th>M8 likelihood</th>
<th>M8a likelihood</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP6</td>
<td>64</td>
<td>1.4</td>
<td>(0.26, 4.9)</td>
<td>-3537.33</td>
<td>-3545.94</td>
<td>0.001</td>
</tr>
<tr>
<td>COX3</td>
<td>44</td>
<td>4.8</td>
<td>(0.26, 4.9)</td>
<td>-3320.11</td>
<td>-3342.84</td>
<td>0.001</td>
</tr>
<tr>
<td>CYTB</td>
<td>96</td>
<td>1.1</td>
<td>(0.26, 4.9)</td>
<td>-5469.81</td>
<td>-5498.22</td>
<td>0.001</td>
</tr>
<tr>
<td>ND5</td>
<td>479</td>
<td>1.9</td>
<td>(0.31, 4.9)</td>
<td>-9664.27</td>
<td>-9693.6</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Position</th>
<th>Ka/Ks</th>
<th>CI</th>
<th>M8 likelihood</th>
<th>M8a likelihood</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP6</td>
<td>64</td>
<td>1.4</td>
<td>(0.26, 4.9)</td>
<td>-3537.33</td>
<td>-3545.94</td>
<td>0.001</td>
</tr>
<tr>
<td>COX3</td>
<td>44</td>
<td>4.8</td>
<td>(0.26, 4.9)</td>
<td>-3320.11</td>
<td>-3342.84</td>
<td>0.001</td>
</tr>
<tr>
<td>CYTB</td>
<td>96</td>
<td>1.1</td>
<td>(0.26, 4.9)</td>
<td>-5469.81</td>
<td>-5498.22</td>
<td>0.001</td>
</tr>
<tr>
<td>ND5</td>
<td>479</td>
<td>1.9</td>
<td>(0.31, 4.9)</td>
<td>-9664.27</td>
<td>-9693.6</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**Table 2. Genes and sites inferred to be under positive selection by Selecton.** Ka/Ks ratio indicates the ratio of nonsynonymous/synonymous mutation rates, CI is the confidence interval of the Ka/Ks ratio defined by the 5th and 95th percentiles of the posterior distribution inferred for the position. The p value is the significance of the likelihood ratio test between the M8 model (selection allowed) vs. M8a model (selection not allowed) of the gene.

**Grant information**

This study was supported by a Faculty Research Grant funded by the University of West Georgia.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Acknowledgements**

Special thanks to Rachel Moran for access to the *E. spectabile* mitogenome sequence, to Frank Fontanella for guidance in the use of Mr. Bayes software, and the Warm Springs Fish Technology Center for darter specimens.
Open Peer Review

Current Peer Review Status: ×

Version 1

Reviewer Report 09 May 2019

https://doi.org/10.5256/f1000research.19195.r47213

© 2019 Kordis D. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dusan Kordis
Department of Molecular and Biomedical Sciences, Jožef Stefan Institute, Ljubljana, 1000, Slovenia

The authors tried to find the evidence of positive selection in 13 mitochondrial proteins, involved in oxidative phosphorylation, in 11 species of darters. Darters are the second largest family of North American fishes and only the minnows (Cyprinidae) have more species. Kral and Watson have found some evidence for positive selection acting on 6 mitochondrial genes, by at least one method used (M8 in Selecton, FEL and MEME in Datamonkey). However, they found that only CYTB gene has evolved under positive selection by all three methods used. The authors expect that a more extensive analysis of mitochondrial gene evolution in all extant darter species may identify lineage specific adaptations and their contributions to the relevant life history traits.

I found a number of problems in the present version of the manuscript:

1. Biology and some special adaptations of darters should be explained in the Introduction. Are there any differences between darter species regarding their life styles, ecology, physiology, metabolism etc.? This information can be used in the interpretation of the observed positive selection in some mitochondrial genes.

2. The authors have just listed the codon positions that have been found to be under positive selection. From the codon numbers, it is very difficult to see what the impact of positively selected sites is. It could be actually biologically important or meaningless.

3. Therefore, the authors must use 3D models of the analysed mitochondrial proteins to show the consequences of the positively selected sites (PSS)/amino acids on the protein secondary and tertiary structure. These amino acids may be functionally important or neutral. The authors should follow some good examples of such analyses (e.g. da Fonseca et al., 20081, Castoe et al., 20082).

4. The authors should indicate the chemical nature of amino acids that are encoded by positively selected codons. Some particular polar amino acids are highly mutable and may not be associated with positive selection (e.g. Xia and Kumar, 20063).
5. The authors should compare their results (e.g. PSS) with the data from literature (e.g. from fishes and other vertebrates) to demonstrate that some positions are conserved PSS or are unique/different PSS.

6. Discussion is very poor and should be strongly improved by obtaining new data about the locations of PSS on the 3D models of the analysed mitochondrial genes. In such a way, their data can be put into biological perspective.

7. The title is uncompelling, merely indicating that the authors haven't obtained any serious conclusions.

References

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
No

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** evolutionary genomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com