Complete mitochondrial genomes of three Neotropical sleeper gobies: *Eleotris amblyopsis*, *E. picta* and *Hemieleotris latifasciata* (Gobiiformes: Eleotridae)

Fernando Alda\textsuperscript{a,b} \textsuperscript{à}, Alexandria J. Adams\textsuperscript{a}, W. Owen McMillan\textsuperscript{b} \textsuperscript{à} and Prosanta Chakrabarty\textsuperscript{a} \textsuperscript{à}

\textsuperscript{a}Museum of Natural Science, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA; \textsuperscript{b}Smithsonian Tropical Research Institute, Panama, Republic of Panama

**ABSTRACT**

We report the first complete mitochondrial genomes of three species of eleotrid fishes from the Pacific and Atlantic watersheds of Panama: *Eleotris amblyopsis*, *E. picta*, and *Hemieleotris latifasciata*. The three species have similar mitochondrial genomes with identical gene order; however, there are differences in the length of control region, 16S rRNA, and in seven of the tRNAs. In addition, ATP8 is one codon shorter in *E. picta* than in *E. amblyopsis* or *H. latifasciata*. We infer a phylogeny for Gobiiformes based on all mitochondrial protein-coding genes, which supports the monophyly of Eleotridae but does not recover Neotropical members of *Eleotris* as a distinct clade.

**Introduction**

The freshwater fish fauna of Middle America contrasts to that of the Nearctic and South American regions in having a large number of salt tolerant species and marine taxa that have invaded the mainland, and are more or less permanent residents of freshwater streams (Myers 1966; Matamoros et al. 2015). This marine-derived and diadromous component comprises five families (Ariidae, Atherinidae, Gerridae, Eleotridae, and Gobiidae) that, despite representing an important fraction of the diversity of this Neotropical assemblage, is considerably understudied (McMahan et al. 2013; Galván-Quesada et al. 2016); particularly relative to the other taxa in the region such as poeciliids and cichlids that have been the focus of many systematic and biogeographic studies (e.g. Martin and Bermingham 1998; Perdices et al. 2005; Alda et al. 2013; Marchio and Piller 2013; McMahan et al. 2017).

In this study, we present the first complete mitochondrial genomes of three fish species of the family Eleotridae that are distributed across Atlantic and Pacific watersheds primarily in the Neotropics (but can be found as far north as California and South Carolina). In addition, we present a mitogenomic phylogeny of the Gobiiformes, which illustrates the position of the Neotropical members of the eleotrids in relation to other species for which complete mitochondrial genomes are available.

**Materials and methods**

As part of ongoing studies on the systematics and biogeography of Neotropical freshwater fishes, we generated full mitogenomes of four individuals from three species of eleotrid fishes: *Eleotris amblyopsis* (STRI-16837, Quebrada Jobito, Río Indio, Coclé, Panama), *E. picta* (STRI-17754, Río Pichende, Río Piña, Darién, Panama; STRI-6991, Río Cate, Veraguas, Panama), and *Hemieleotris latifasciata* (STRI-7098, Río Tebario, Veraguas, Panama). Specimens were collected by electrofishing or seining, and Gill arches were excised and preserved in a saturated 20% dimethyl sulphoxide (DMSO) and 0.5 M EDTA pH 8 solution at 4°C. The whole specimens were fixed in formalin, vouchered and deposited in the Neotropical Fish Collection (NFC-STRI) at the Smithsonian Tropical Research Institute in Panama (STRI).

Mitogenome sequences were obtained as a by-product of a hybrid target capture of ultraconserved elements (UCEs) (Faircloth et al. 2012). We trimmed our reads for low-quality bases and adapter contamination using the parallel wrapper around Trimmomatic (Bolger et al. 2014) in illumiprocessor (Faircloth 2013). Then, we mapped our reads to reference mitogenomes from the evolutionary closest species available using Bowtie 2 (Langmead and Salzberg 2012) implemented in Geneious 10.1.3 (www.geneious.com; Kearse et al. 2012). Using Geneious 10.1.3, we then created consensus sequences and annotated them for each sample.

We aligned our mitogenomes with representatives from five of the seven families of Gobiiformes (sensu Agorreta et al. 2013), and extracted all 13 protein-coding genes for subsequent phylogenetic analysis. We partitioned the data by gene and by codon, and estimated the best arrangement of partitions under a GTR+G nucleotide substitution model using PartitionFinder (Lanfear et al. 2012). We used these partitions in a Maximum Likelihood analysis in RAxML (Stamatakis 2014), that we run for 40 searches to find the
Table 1. Characteristics of the mitochondrial genomes of Eoleotris amblyopsis, E. picta, and Hemileotris latifasciata.

| Code | Amino acid/gene | Start | Stop | Size | Spacer (+) or overlap (-) | Direction | Start codon | Stop codon |
|------|----------------|-------|------|------|---------------------------|-----------|-------------|------------|
| F    | tRNA-Phe       | 1/1/1 |      | 68/68/68 |                           |           |             |            |
| V    | tRNA-Thr       | 1022/1022/1022 |      | 72/72/71 |                           |           |             |            |
| L    | tRNA-Leu       | 2780/2780/2782 |      | 2853/2851/2856 |                       | 74/74/75 |             |            |
| I    | tRNA-Ile       | 3834/3841/3836 |      | 3903/3910/3905 |                   |            | 70/70/70   |            |
| Q    | tRNA-Gln       | 3973/3980/3975 |      | 3903/3910/3905 |               | 71/71/71   |             |            |
| M    | tRNA-Met       | 3973/3980/3975 |      | 4042/4049/4043 |              | 70/70/70   |             |            |
| W    | tRNA-Trp       | 4043/4050/4044 |      | 5089/5096/5090 |            | 1047/1047/1047 | 1+/1+/1- |            |
| A    | tRNA-Ala       | 5243/5239/5233 |      | 5166/5171/5165 |               | 69/69/69   |             |            |
| N    | tRNA-Asn       | 5308/5313/5307 |      | 5236/5241/5235 |            | 73/73/72   |             |            |
| C    | tRNA-Cys       | 5410/5416/5410 |      | 5345/5351/5345 |               | 66/66/66   |             |            |
| Y    | tRNA-Tyr       | 5480/5487/5481 |      | 5411/5417/5411 |                   | 70/71/71   |             |            |
| S    | tRNA-Ser       | 7106/7113/7107 |      | 7036/7043/7037 |            | 71/71/71   |             |            |
| D    | tRNA-Asp       | 7110/7117/7111 |      | 7181/7189/7182 |            | 72/73/72   |             |            |
| K    | tRNA-Lys       | 7879/7886/7880 |      | 7953/7960/7954 |               | 75/75/75   |             |            |
| ATP8 |                | 8735/8742/8740 |      | 8739/8743/8740 |            | 1089/1089/1089 | 1+/1+/1- |            |
| COXI |                | 5482/5489/5483 |      | 7035/7042/7036 |               | 154/154/154 |             |            |
| S    | tRNA-Ser       | 7106/7113/7107 |      | 7036/7043/7037 |            | 71/71/71   |             |            |
| D    | tRNA-Asp       | 7110/7117/7111 |      | 7181/7189/7182 |            | 72/73/72   |             |            |
| K    | tRNA-Lys       | 7879/7886/7880 |      | 7953/7960/7954 |               | 75/75/75   |             |            |
| ATP8 |                | 8735/8742/8740 |      | 8739/8743/8740 |            | 1089/1089/1089 | 1+/1+/1- |            |
| COXI |                | 5482/5489/5483 |      | 7035/7042/7036 |               | 154/154/154 |             |            |
| S    | tRNA-Ser       | 7106/7113/7107 |      | 7036/7043/7037 |            | 71/71/71   |             |            |
| L    | tRNA-Leu       | 1180/1189/11888 |     | 1196/1196/11960 |          | 73/73/73 |            |            |
| NDS  |                | 11963/11967/11961 |      | 13801/13805/13799 |           | 1839/1839/1839 | 4+/4+/4- |            |
| E    | tRNA-Glu       | 14369/14392/14386 |       | 14320/14324/14318 |        | 67/69/69 |             |            |
| T    | tRNA-Thr       | 15354/15452/15533 |      | 15606/15615/15604 |          | 73/74/72 |             |            |
| P    | tRNA-Pro       | 15677/15687/15674 |      | 15680/15617/15605 |          | 70/71/70 |             |            |
| Control |          | 15678/15686/15675 |      | 16519/16547/16523 |          | 842/860/849 |             |            |

best tree, and performed 500 bootstrap replicates to assess nodal support.

Results and discussion

We obtained complete mitochondrial genome sequences from three representatives of the Eleotridae found primarily in the Neotropical: *E. amblyopsis, E. picta,* and *H. latifasciata.* All species genomes consist of 22 tRNA genes, two rRNA genes, 13 protein-coding genes, and control region (D-loop), arranged in identical order; their gene lengths are 16,519 bp, 16,547 bp, and 16,523 bp, respectively (Table 1). Full sequences are deposited in GenBank under accession numbers: MF927490 (*E. amblyopsis* STRI-16837), MF927491 (*E. picta* STRI-6991), MF927492 (*E. picta* STRI-17754), and MF927495 (*H. latifasciata* STRI-7098).

The overall base composition is 29.2% A, 48.9% C, 16.3% G, 25.5% T, and 45.2% GC for *E. amblyopsis.* For *E. picta,* the base composition is 27.6% A, 30.9% C, 17.7% G, 23.8% T, and 48.6% GC. For *H. latifasciata,* the overall base composition is 31% A, 36.3% C, 10.1% G, 22.6% T, and 46.4% GC.

All protein-coding genes are of identical length across the three species, with the exception of ATPB that is reduced by a single codon in *E. picta.* Non-coding regions, on the other hand, show more variation in length. For example, seven out of 22 tRNAs differ in size by one single base pair (bp): *E. amblyopsis* is one bp longer in tRNA-Trp, and one bp shorter in tRNA-Tyr; *E. picta* has an additional bp in tRNA-Pro and tRNA-Asp; *H. latifasciata* is one bp shorter in tRNA-Val, tRNA-Met, tRNA-Asn, and tRNA-Thr. 16S rRNA and control region differ drastically among the three species, ranging between 1687 and 1695 bp for the 16S rRNA, and 842 and 860 b for the control region. Within species, the two *E. picta* specimens exhibit 58 mutational differences throughout their mitochondrial genomes; 15 of these changes are found in non-coding regions, while the translation of eight codons differs as a result of these mutations.

The most common initiation codon, ATG, is found in all the mitochondrial protein-coding genes apart from COXI, that has GTG (Val), and ND6, that has TAC (Tyr), as start codons in all species. Five genes share the common termination codon TAA across all species. In ND1, *E. amblyopsis* and *E. picta* differ in having the termination codon TAG, while in *E. picta* remains TAA. For ND6, ATC is the stop codon in every species. COXII, COXIII, ND3, ND4, and CytB coding regions have incomplete stop codons: T or TA. These incomplete stop codons are completed as TAA by the post-transcriptional polyadenylation of the corresponding mRNAs (Ojala et al. 1981).

Our phylogenetic reconstruction strongly supports the monophyly of the family Eleotridae. *Hemileotris latifasciata* is the sister group to a clade containing all the species of the
genus *Eleotris*. Neotropical members of *Eleotris* do not form a monophyletic group: *E. picta*, from the Eastern Pacific, is the sister species to the clade formed by *E. amblyopsis*, distributed in the East Atlantic, and *E. acanthopoma*, which is widespread across the South West Pacific. *Eleotris oxycephala* from the West Pacific, is the sister taxon to the least inclusive clade that contains all the Neotropical members of *Eleotris* (Figure 1). We resist drawing any major biogeographic conclusion at this time, given the limited sampling, but do point out that major transoceanic relationships do exist in Gobiiformes (Chakrabarty et al. 2012).

Here, we present the first mitogenomes of three members of the Eleotridae from the Neotropics. These sequences will aid future studies into the evolutionary relationships of gobiform fishes. This work will also aid our understanding of the diversity of Middle America freshwater fish assemblages as mitochondrial genomes of additional species and lineages become available.

**Acknowledgements**

We thank Ruth G. Reina for laboratory and curatorial assistance. We also acknowledge Brant Faircloth for leading the workshop during which these data were generated.

**Disclosure statement**

The authors declare no conflict of interest.

**Funding**

This work was funded by the Smithsonian Institution and the Smithsonian Tropical Research Institute Office of Fellowships, and a National Science Foundation Grant to PC [grant number DEB-1354149].

**ORCID**

Fernando Alda [http://orcid.org/0000-0002-6858-9566](http://orcid.org/0000-0002-6858-9566)
W. Owen McMillan [http://orcid.org/0000-0003-2805-2745](http://orcid.org/0000-0003-2805-2745)
Prosanta Chakrabarty [http://orcid.org/0000-0003-0565-0312](http://orcid.org/0000-0003-0565-0312)

**References**

Agorreta A, San Mauro D, Schliewen U, Van Tassell JL, Kovacic M, Zardoya R, Rüber L. 2013. Molecular phylogenetics of Gobioidei and phylogenetic placement of European gobies. Mol Phylogenet Evol. 69:619–633.
Alda F, Reina RG, Doadrio I, Bermingham E. 2013. Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. Mol Phylogenet Evol. 66:1011–1026.
Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Chakrabarty P, Davis MP, Sparks JS. 2012. The first record of a transoceanic sister-group relationship between obligate vertebrate troglobites. PLoS One. 7:e44083.

Faircloth BC. 2013. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. http://dx.doi.org/10.6079/J9ILL.

Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst Biol. 61:717–726.

Galván-Quesada S, Doadrio I, Alda F, Perdices A, Reina RG, García Varela M, Hernández N, Mendoza AC, Bermingham E, Domínguez-Domínguez O. 2016. Molecular phylogeny and biogeography of the amphidromous fish genus Dormitator Gill 1861 (Teleostei: Eleotridae). PLoS One. 11:e0153538.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28:1647–1649.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 29:1695–1701.

Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9:357–359.

Marchio EA, Piller KR. 2013. Cryptic diversity in a widespread live-bearing fish (Poeciliidae: Belonesox). Biol J Linn Soc. 109:848–860.

Martin AP, Bermingham E. 1998. Systematics and evolution of lower Central American cichlids inferred from analysis of cytochrome b gene sequences. Mol Phylogenet Evol. 9:192–203.

Matamoros WA, McMahan CD, Chakrabarty P, Albert JS. 2015. Derivation of the freshwater fish fauna of Central America revisited: Myer’s hypothesis in the twenty-first century. Cladistics. 31:177–188.

McMahan CD, Davis MP, Domínguez-Domínguez O, García-de-León FJ, Doadrio I, Piller KR. 2013. From the mountains to the sea: phylogeography and cryptic diversity within the mountain mullet, Agonostomus monticola (Teleostei: Mugilidae). J Biogeogr. 40:894–904.

McMahan CD, Ginger L, Cage M, David KT, Chakrabarty P, Johnston M, Matamoros WA. 2017. Pleistocene to holocene expansion of the black-belt cichlid in Central America, Vieja maculicauda (Teleostei: Cichlidae). PLoS One. 12:e0178439.

Myers GS. 1966. Derivation of the freshwater fish fauna of Central America. Copeia. 1966:766–773.

Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature. 290:470–474.

Perdices A, Doadrio I, Bermingham E. 2005. Evolutionary history of synbranchid eels (Teleostei: Synbranchiformes) in Central America and the Caribbean islands inferred from their molecular phylogeny. Mol Phylogenet Evol. 37:460–473.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.