A Unique Mitochondrial Gene Block Inversion in Antarctic Trematomin Fishes: A Cautionary Tale

Selina Patel, Clive W. Evans, Alex Stuckey, Nicholas J. Matzke, and Craig D. Millar

From the School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand (Patel, Evans, Matzke, and Millar); and Genomics England, Queen Mary University of London, Dawson Hall, London EC1M 6BQ, UK (Stuckey).

Address correspondence to C. D. Millar at the address above, or e-mail: cd.millar@auckland.ac.nz.

Corresponding Editor: C. Scott Baker

Abstract

Many Antarctic notothenioid fishes have major rearrangements in their mitochondrial (mt) genomes. Here, we report the complete mt genomes of 3 trematomin notothenioids: the bald notothen (Trematomus (Pagothenia) borchgrevinki), the spotted notothen (T. nicolai), and the emerald notothen (T. bernacchii). The 3 mt genomes were sequenced using next-generation Illumina technology, and the assemblies verified by Sanger sequencing. When compared with the canonical mt gene order of the Antarctic silverfish (Pleuragramma antarctica), we found a large gene inversion in the 3 trematomin mt genomes that included tRNA\text{Ile}, ND1, tRNA\text{His}, ND5, ND6, tRNA\text{Glu}, tRNA\text{Gln}, tRNA\text{Val}, 12S, tRNA\text{Tyr}, and the control region. The trematomin mt genomes contained 3 intergenic spacers, which are thought to be the remnants of previous gene and control region duplications. All control regions included the characteristic conserved regulatory sequence motifs. Although short-read next-generation DNA sequencing technology has allowed the rapid and cost-effective sequencing of a large number of complete mt genomes, it is essential in all cases to verify the assembly in order to prevent the publication and use of erroneous data.

Key words: mitochondrial genome, Notothenioidei, Trematomus bernacchii, Trematomus (Pagothenia) borchgrevinki, Trematomus nicolai

The teleost fauna of the Southern Ocean is dominated by a clade of periform fishes belonging to the suborder Notothenioidei (Eastman 1993). This suborder is comprised of 8 families, of which the 3 basal lineages (Bovichtidae, Pseudaphritidae, and Eleginopsidae) are predominantly non-Antarctic in distribution. Most species in the remaining 5 families (Nototheniidae, Harpagiferidae, Artedidraconidae, Bathydraconidae, and Channichthyidae) are endemic to the Southern Ocean (Eastman 2005).

The phylogenetic relationships of the notothenioids have been studied for more than 2 decades. Recently, the phylogeny of this group has been reanalyzed and updated using a combination of mitochondrial (mt) and nuclear DNA markers, including 15 new complete and 2 partial mt genomes (Papetti et al. 2021). The study identified a number of novel rearrangements in the mt genomes, including an extremely rare inversion event in the Trematominae. Papetti et al. (2021) generated a new phylogeny which showed that the mt evolution of the notothenioids has been characterized by multiple, relatively rapid changes in mt gene order.

With the advent of next-generation DNA sequencing, complete mt genomes have become much more cost effective and feasible to collect. However, the assembly of these genomes often relies on using a scaffold from a closely related species. This assumes the gene order between these genomes is the same or very similar, and ignores the possibility of major gene rearrangements. Furthermore, mt genomes obtained from the assembly of short reads from next-generation DNA sequencing are seldom verified by targeted PCR amplification, potentially resulting in the publication of incorrectly assembled genomes.

The vertebrate mt genome is highly conserved, consisting of 13 protein-coding genes, 2 ribosomal RNA (rRNA) genes, and 22 transfer RNA (tRNA) genes (Wolstenholme 1992). In addition, there are a number of noncoding regions, the most significant of which is the control region (CR) that contains transcriptional promoters for both the heavy (H) and light (L) strands. Apart from the presence of the D-loop, the CR can be divided into 3 domains: a domain associated with the termination-associated sequences, a conserved central domain (CCD), and conserved sequence block (CSB) domains (Anderson et al. 1981; Brown et al. 1986). The order of the 37 genes and noncoding regions in the mt genome tends to be conserved among most vertebrate species studied to date, although deviations from the canonical order have been identified in various groups, including fishes (Satoh et al. 2016).

Changes in gene order can be misinterpreted as gene loss unless detailed analyses are undertaken, as shown in the supposed loss of the mt NADH dehydrogenase subunit 6 (ND6) and tRNA\text{Tyr} genes in notothenioids (Papetti et al. 2007). In fact, in the 5 Antarctic notothenioid families studied, these
2 genes were not lost but simply translocated from their canonical location between the NADH dehydrogenase subunit 5 (ND5) and cytochrome b (Cytb) genes to the CR, and subsequently overlooked (Zhuang and Cheng 2010). Papetti et al. (2021) have recently shown that whole mt genomes of Antarctic notothenioids vary greatly with respect to gene order. Their study was the first to report novel gene orders from representative species of all Nototheniidae families, including the trematomins.

In the present study, we identified a unique gene order when assembling the complete mt genomes from sequences of 3 Antarctic fish species: the bald notothen (Trematomus (Pagothenia) borchgrevinki), the spotted notothen (T. nicolai), and the emerald notothen (T. bernacchii). The gene order was verified using PCR and Sanger sequencing. We also characterized the CR domains in these 3 species.

**Methods**

**Sample Collection and DNA Extraction**

Adult specimens of the 3 trematomins (T. borchgrevinki, T. nicolai, and T. bernacchii) were collected from the vicinity of Ross Island, McMurdo Sound, Antarctica using routine fishing methods. Each specimen was identified using morphological features and specimen identification was confirmed using standard DNA barcoding methods (Ratnasingham and Hebert 2007). Additional specimen details including collection date and location can be found in Table 1.

Cells were scraped from the gill tissue and re-suspended in STE (50 mM NaCl, 50 mM Tris-HCl, 100 mM EDTA, pH 8.0) buffer. Genomic DNA was isolated from the suspension using Proteinase K digestion followed by a phenol: chloroform: isoamyl extraction (Sambrook et al. 1989). DNA was then digested with 20 μg/μL RNase at 37 °C for 4 h.

**Library Construction and Assembly**

Libraries were constructed using the Illumina TruSeq Nano kit (2 × 250 bp reads) or the Affymetrix Prep2Seq kit (2 × 300 bp reads) and sequenced on the Illumina MiSeq (San Diego, CA) platform. For T. borchgrevinki, an additional library was constructed using Rubicon Thruplex DNA-seq (2 × 125 bp paired-end) and sequenced on the Illumina HiSeq 2500 (San Diego, CA) platform using v4 chemistry.

Sequencing data were quality checked with FastQC (Andrews 2010) to ensure that there were no issues with the sequencing process or the resulting data. The data were then mapped against 2 reference genomes: the black notothen (Notothenia coriiceps) full nuclear genome and the N. coriiceps mt genome (accession numbers: AZAD00000000, NC_015653) using the tool bbmap package to isolate fish-only reads (Bushnell 2014). To map to a reference, a candidate sequence needed to share at least 97% identity. The mapping was repeated with a minimum of 90% identity, using the output of the previous step. Raw reads were then mapped against the mt genomes of T. borchgrevinki (accession number KU951144.1) and 3 closely related species: the Patagonian toothfish (Dissostichus eleginoides), N. coriiceps, and the Antarctic silverfish (Pagothenia antarctica) (accession numbers NC_018135.1, NC_015653, and JF933905, respectively) in order to extract only the mt DNA reads. The mt DNA sequence pools were assembled into one contig which, when annotated with the MITOS web server (Bernt et al. 2013), contained the COI gene. This gene was used as a seed for MITObim (Hahn et al. 2013) to assemble the mt genomes. Each of the assembled mt genomes of T. borchgrevinki, T. bernacchii, and T. nicolai showed a large gene block inversion.

To verify the inversion, a ~9 kb fragment of the mt genome was amplified using primers designed to Cytb (forward primer) and ND2 (reverse primer). A long-range PCR was carried out using the TaKaRa LA Taq kit according to the manufacturer’s instructions. Once the gene orientation was verified at either ends of this ~9 kb fragment through Sanger sequencing, smaller overlapping fragments were then amplified for the remaining region (see Supplementary Table S1 for primer details). PCRs were run in 25 μL volume reactions and included 10–30 ng template DNA, 2 mM MgCl2, 0.4 μM forward and reverse primers, and 0.1 U Taq (Life Technologies). In some instances, betaine and dimethyl sulfoxide (DMSO) were used in both the PCR and cycle sequencing reactions for areas containing repetitive sequences to inhibit secondary structure formation. The thermal cycling conditions were as follows: 2 min at 94 °C followed by 30 cycles of 30 s at 94 °C, 45 s at the annealing temperature (Supplementary Table S1) and 60 s at 72 °C, and a final extension of 5 min at 72 °C.

All PCR products were purified, cycle sequenced using Big Dye 3.1 chemistry, and subsequently analyzed on an ABI Prism 3130xl genetic analyzer (Applied Biosystems). PCR products were sequenced in both directions. Sequences were edited manually using GENEIOUS (http://www.geneious.com/) and then aligned to the assembled complete mt genomes. The mt genomes of all 3 species had a ~9 kb fragment verified by PCR. An exception to this was a ~500 bp region of the T. bernacchii CR which proved difficult to amplify due to the presence of repetitive sequences.

**Results and Discussion**

**Illumina Sequencing**

The Illumina TruSeq Nano libraries produced 33 million barcoded reads for each of the 3 samples (T. borchgrevinki, T. nicolai, and T. bernacchii). The Affymetrix Prep2Seq libraries

| Table 1. Sample identifiers and collection details for the bald notothen (Trematomus borchgrevinki), spotted notothen (T. nicolai), and emerald notothen (T. bernacchii) |
| Species                  | Sample ID | Collection date | Collection coordinates          |
|--------------------------|-----------|-----------------|---------------------------------|
| T. borchgrevinki         | 11/134    | December 2011   | Between 77.635 and 77.885°S and 166.311 and 166.770°E |
| T. nicolai               | 11-10     | December 2011   |                                  |
| T. bernacchii            | 11/145    | December 2011   |                                  |
produced 28.8 million barcoded reads for *T. borchgrevinki*, 28.8 million barcoded reads for *T. nicolai*, and 14.4 million barcoded reads for *T. bernacchii*. The *T. borchgrevinki* library from the Rubicon Thruplex DNA-seq kit produced 440 million barcoded paired-end reads.

**Mitochondrial Genome Organization**

The lengths of the mt genomes of the 3 trematomin species under study were 18,981 bp (*T. borchgrevinki*), 19,358 bp (*T. bernacchii*) (Figure 1). They are registered in GenBank under accession numbers MZ779011, MZ779013, and MZ779012, respectively. Each mt genome contained 13 protein-coding genes, 22 tRNA genes, 1 large intergenic spacer (UN3), and 19,795 bp (*T. bernacchii*)), 19,358 bp (*T. borchgrevinki*), and 18,981 bp (*T. nicolai*), respectively. Each mt genome was 24% A, 31.6% T, 21.9% G, and 22% C.

A number of studies have published the complete mt genomes of trematomin species (*T. loennbergii*, *T. borchgrevinki*, *T. bernacchii*, and *T. pennellii*), but surprisingly none has reported any differences in the canonical gene order (Liu et al. 2016; Song et al. 2016; Alam et al. 2019; Choi et al. 2021). Additionally, Song et al. (2016) published the *T. bernacchii* mt genome with an incomplete ND6 and an incomplete 12S sequence. Furthermore, 2 of the above studies do not provide any information on how the sequence data were obtained, their methods of assembly or data verification. The GenBank submission for these 2 studies also lacks any accompanying metadata. In order to evaluate the assemblies of complete mt genomes, we strongly recommend that the sequencing data be made available to other researchers on request or deposited in an appropriate database.

Based on our results and the recent findings of Papetti et al. (2021), we believe that the gene order reported by Liu et al. (2016), Song et al. (2016), Alam et al. (2019), and Choi et al. (2021) is incorrect. Furthermore, due to the lack of information regarding the methods used, it is very difficult to identify the source(s) of the likely problems, for example, incorrect assembly; the use of closely related mt genomes as a scaffold; and reliance on only short read next-generation sequencing. As a result, these mt genome assemblies should not be used in any future analysis or at the very least viewed with caution.

Importantly, it is now possible to overcome many of the problems discussed above and the issue of inclusion of numts in the mt genome assemblies. Although, Illumina sequencing is readily available and cost effective, it generates large numbers of short sequences (100–300 bp), which require assembly. In contrast, long-read sequencing technology, such as PacBio and Nanopore, has the potential to sequence the entire mt genome in a single read. PacBio is relatively expensive and is therefore unlikely to be used to sequence novel mt genomes. However, Nanopore technology in combination with long-range PCR and pooling of individually barcoded samples allows fast and cost-effective sequencing of entire mt genomes (Formenti et al. 2021).

**Mitochondrial Gene Rearrangements**

We found a major rearrangement in the gene order between the 3 trematomin species and that of other notothenioids such as *P. antarctica* and *N. coriiceps*. The most significant differences observed were a large gene block inversion of the mt genome region which falls between tRNA\(^{Glu}\) and tRNA\(^{Glu}\) of the *P. antarctica* mt genome, and the presence of 3 intergenic spacers. The gene block inversion contained 7 genes and 2 noncoding regions in the following order; an intergenic spacer (UN3), followed by tRNA\(^{Gln}\), ND1, tRNA\(^{Lys}\), 16S, tRNA\(^{Val}\), 12S, tRNA\(^{Phe}\), and ended with the CR. All 3 trematomins showed the same general gene order and pattern except for the intergenic spacers and CRs being of differing lengths (Figure 1a–c; Table 2).

The gene complements of the 3 trematomin mt genomes reported here are the same as the basal non-Antarctic Bovichtus species (Satoh et al. 2016). However, the gene order is notably different between ND5 and tRNA\(^{Glu}\) (Figure 1). According to Zhuang and Cheng (2010), a tandem gene duplication event occurred between the basal non-Antarctic bovichtids and the common ancestor of the Antarctic clade. This duplication was followed by the early loss/degredation of ND6, tRNA\(^{Glu}\), and Cytb leading to “Pattern I” as described by these authors and exemplified in the mt genome of the extant *P. antarctica*. For the trematomins described in this paper, a possible evolutionary pathway from this point to the current trematomin gene order has been detailed by Papetti et al. (2021), who now refer to this pathway as “TremaGo.” This involves 1) partial random loss of CR1 in the *P. antarctica* mt genome as evidenced by the lack of characteristic CR conserved sequences (e.g., extended termination-associated [ETASs] and CSBs); 2) partial random loss of tRNA\(^{Phe}\) and tRNA\(^{Phe}\) between ND6 and CR2; and 3) inversion of the gene block CR2, tRNA\(^{Phe}\), 12S, tRNA\(^{Val}\), 16S, tRNA\(^{Lys}\), ND1, and tRNA\(^{Phe}\). It is important to note that intergenic spacers are generated during this pathway which are thought to include the remnants of lost/duplicated genes. Overall, the results of the mt genome assembly of *T. borchgrevinki* are consistent with those shown by Papetti et al. (2021).

**Control Region**

The CR of mammalian mtDNA typically lies between the tRNA\(^{Pro}\) and tRNA\(^{Phe}\) genes, and this is reflected in the mt genome of the basal non-Antarctic notothenioid thornfish (*Bovichtus argentinus*) (Satoh et al. 2016). In the trematomin studied here, the position of the tRNA\(^{Phe}\) gene which delineates one end of the canonical mammalian CR is now occupied by tRNA\(^{Glu}\) due to a gene block inversion.

Comparative sequence analysis of the CR between the 3 trematomins and other notothenioids allowed us to infer the presence of 2 extended termination-associated sequences (ETAS1 and ETAS2) within the CR domain, each 31 nucleotides long (Figure 2). These are significantly shorter than the ~60 bp sequences originally identified by Sbisà et al. (1997), but are largely consistent with the sequences...
identified by Zhuang and Cheng (2010). Both ETAS domains in all 3 trematomins contain the sequence 5'-ATGA-3' (with reference to the L-strand) as the complementary termination-associated sequence (cTAS). This sequence is at the 5' end of a 15 bp sequence referred to as the coreTAS on the L-strand in humans (Jemt et al. 2015).

In most marine teleosts, the CCD contains 3 CSBs (CSB-F, CSB-E, and CSB-D), and we confirmed their presence in
the CRs of the trematomins studied (Figure 2). CSB-F is positioned nearest the ETAS region and has a cTAS sequence at its 5′ end. The GTGGG box identified in the CSB-E domain of many teleost species was present as a modified GTGAG sequence in the trematomins. The CSB-D sequence was found within the CCD region as identified in groupers with minor variation (Zhuang et al. 2013). We also identified 2 conserved sequences in the CR CSB domain (CSB-1 and CSB-2). The CSB-1 region had the characteristic CATAA sequence at its 3′ end, while CSB-2 had the characteristic poly C stretch separated by TA (Zhuang et al. 2013).

A comparison between the mt genome sequences of *T. borchgrevinki* from our study and that of Papetti et al. (2021) nonetheless revealed several differences. The genome of *T. borchgrevinki* in our study was found to be a total of 656 bp longer. Much of the difference in length (644 bp) was found to be within the CR. This CR length difference could be further broken down into 3 regions, which contained extra

| Gene   | Abbreviation | T. borchgrevinki | T. nicolai | T. bernacchii |
|--------|--------------|-----------------|------------|-------------|
| ND4L   |              | 1 297 297       | 1 297 297  | 1 297 297   |
| ND4    |              | 291 1671 1381   | 291 1671 1381 | 291 1671 1381 |
| tRNA^Hts| H            | 1672 1740 69    | 1672 1740 69 | 1672 1740 69 |
| tRNA^Ser| S1           | 1741 1807 67    | 1741 1807 67 | 1741 1807 67 |
| tRNA^Ess| L1           | 1812 1884 73    | 1812 1884 73 | 1812 1884 73 |
| ND5    |              | 1885 3723 1839  | 1885 3723 1839 | 1885 3723 1839 |
| Intergenic spacer | UN1 | 3724 3765 43 | 3724 3771 49 | 3724 3770 48 |
| Cyt b  |              | 3766 4906 1141  | 3772 4912 1141  | 3771 4911 1141  |
| tRNA^Ths| T            | 4907 4978 72    | 4913 4984 72 | 4912 4983 72 |
| tRNA^Pass| P           | 4978 5047 70    | 4984 5053 70 | 4983 5052 70 |
| Intergenic spacer | UN2 | 5048 5048 338  | 5054 5061 609 | 5053 5062 611 |
| ND6    |              | 5430 5948 519   | 5662 6180 519 | 5663 6181 519 |
| tRNA^Gns| E            | 5949 6016 68    | 6181 6248 68 | 6182 6249 68 |
| Intergenic spacer | UN3 | 6017 6563 547  | 6249 7207 960 | 6250 7368 1120 |
| tRNA^The| I            | 6564 6633 70    | 7208 7277 70 | 7369 7438 70 |
| ND1    |              | 6638 7612 975   | 7282 8256 975 | 7443 8417 975 |
| tRNA^Ess| L2           | 7613 7686 74    | 8257 8330 74 | 8418 8491 74 |
| 16S    |              | 7687 9375 1689  | 8331 10 019 1689 | 8492 10 183 1692 |
| tRNA^Val| V            | 9377 9448 72    | 10 021 10 092 72 | 10 185 10 256 72 |
| 12S    |              | 9452 10 396 945 | 10 096 11 040 945 | 10 260 11 203 944 |
| tRNA^Pas| F            | 10 397 10 464 68 | 11 041 11 108 68 | 11 204 11 271 68 |
| Control region |  | 10 465 12 808 2344 | 11 109 13 187 2079 | 11 272 13 612 2341 |
| tRNA^Gns| Q            | 12 809 12 880 72 | 13 188 13 259 72 | 13 613 13 684 72 |
| tRNA^Mar| M            | 12 880 12 948 69 | 13 259 13 327 69 | 13 684 13 752 69 |
| ND2    |              | 12 949 13 994 1046 | 13 328 14 373 1046 | 13 753 14 798 1046 |
| tRNA^7ps| W            | 13 995 14 065 71 | 14 374 14 444 71 | 14 799 14 869 71 |
| tRNA^Ala| A            | 14 067 14 135 69 | 14 446 14 514 69 | 14 871 14 939 69 |
| tRNA^Asn| N            | 14 137 14 209 73 | 14 516 14 588 73 | 14 941 15 013 73 |
| tRNA^Cys| C            | 14 235 14 300 66 | 14 612 14 677 66 | 15 042 15 107 66 |
| tRNA^Gly| Y            | 14 301 14 371 71 | 14 678 14 748 71 | 15 108 15 178 71 |
| COI    |              | 14 373 15 923 1551 | 14 750 16 300 1551 | 15 180 16 730 1551 |
| tRNA^Ser2| S2           | 15 924 15 994 71 | 16 301 16 371 71 | 16 738 16 808 71 |
| tRNA^Asp| D            | 15 996 16 066 71 | 16 373 16 443 71 | 16 810 16 880 71 |
| COII   |              | 16 069 16 759 691 | 16 446 17 136 691 | 16 883 17 573 691 |
| tRNA^His| K            | 16 760 16 833 74 | 17 137 17 210 74 | 17 574 17 647 74 |
| ATP8   |              | 16 835 17 002 168 | 17 212 17 379 168 | 17 649 17 816 168 |
| ATP6   |              | 16 981 17 676 696 | 17 358 18 053 696 | 17 795 18 490 696 |
| COIII  |              | 17 709 18 493 785 | 18 086 18 870 785 | 18 523 19 307 785 |
| tRNA^Gly| G            | 18 494 18 563 70 | 18 871 18 940 70 | 19 308 19 377 70 |
| ND3    |              | 18 564 18 912 349 | 18 941 19 289 349 | 19 378 19 726 349 |
| tRNA^Arg| R            | 18 913 18 981 69 | 19 290 19 358 69 | 19 727 19 795 69 |

Table 2. Mitogenome organization of the bald notothen (*Trematomus borchgrevinki*), spotted notothen (*T. nicolai*), and emerald notothen (*T. bernacchii*).
sequences of 404 bp, 46 bp, and 194 bp, respectively. Given that no differences were found in any of the coding genes, we believe these CR differences to be real rather than sequencing/assembly errors, and possibly the result of slippage replication of the repetitive regions in the CR. If verified, the CR variation among individuals of *T. borchgrevinki* and potentially other trematomins is an important novel finding.

With next-generation sequencing and indexing technologies, it is now practical to sequence the CR from a large number of individuals of a single species in order to investigate if population-level genetic variation is widespread. Complete CR sequences from multiple individuals and samples collected from different geographic regions of *T. borchgrevinki* would be necessary to establish the nature and extent of the CR variation. CR sequences are known to vary greatly within and between species, and have been used as valuable population genetic markers (Avise 2004; Jamandre et al. 2014). Based on the CR differences reported here, this may also be the case in *T. borchgrevinki* and other trematomins.

Complete mt genome data from notothenioids has typically been used for phylogenetic analyses. As a result there is, in many cases, complete mt genomic data available from only a single individual or at best a very small number of individuals for each species. In contrast, Lin et al. (2012) sequenced large regions of the mt genomes of the mackerel icefish (*Champsocephalus gunnari*) from 32 individuals and interestingly found variation in the number of CRs and genes among individuals. Similar studies of the mt genomes from *T. borchgrevinki* and other trematomins would be important, in order to establish if this finding occurs in other species within the family.

**Conclusions**

Each of the mt genomes of 3 trematomins species (*T. borchgrevinki*, *T. nicolai*, and *T. bernacchii*) was found to have a large, unique gene block inversion. These results provide evidence that many of the published genome assemblies for *T. borchgrevinki* and other trematomins are incorrect. There is significant length variation in the CR in *T. borchgrevinki* between the 2 individuals compared, one from this study and the other from Papetti et al. (2021). Recent advances in DNA sequencing technology and associated bioinformatic pipelines will lead to large numbers of high-quality, error-free mt genomes in the near future (Formenti et al. 2021). Furthermore, large-scale complete mt sequencing of a range of notothenioids species including the trematomins would provide valuable insights into the population genetics and the evolution of fish mt genomes.

**Supplementary Material**

Supplementary material is available at *Journal of Heredity* online.

**Funding**

This work was supported by the University of Auckland research funds and the Allan Wilson Centre for Molecular Ecology and Evolution.

**Acknowledgments**

We thank Auckland Genomics for DNA sequencing support. We are grateful to colleagues at McMurdo Station and Scott.
Data Availability

The sequence data are available in GenBank under accession numbers MZ779011, MZ779013, and MZ779012 for T. borchgrevinki, T. nicolai, and T. bernacchii, respectively. Raw Illumina data are available on Dryad Digital Repository https://datadryad.org/stash/share/cpEOPUfS6t6lkuIhwhsjKHhn_-j2GU5xpylpFc49vVY (Patel et al. 2022).

References

Alam MJ, Kim J-H, Andriyono S, Lee J-H, Lee SR, Park H, Kim H-W. 2019. Characterization of complete mitochondrial genome and gene organization of sharp-spined notothenia, Trematomus pennellii (Perciformes: Nototheniidae). Mitochondrial DNA B. 4:648–649.

Andrews S, Bankier A, Barrell B, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–465.

Antunes A, Ramos MJ. 2005. Discovery of a large number of previously uncharacterized mitochondrial pseudogenes in fish genomes. Genomics 86:708–717.

Avise JC. 2004. Molecular markers, natural history and evolution. 2nd ed. Sunderland, (MA): Sinauer Associates.

Bernt M, Donath A, Jühling F, Externbrink F, Drouin J, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69:313–319.

Brown GG, Gadaleta G, Pepe G, Saccone C, Sbisà E. 1986. Structural and genetic conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. J Mol Biol. 192:503–511.

Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. Lawrence Berkeley National Laboratory. LBNL Report #: LBNL-7065E. Available from: https://esescholarship.org/uc/item/1h3515gn

Choi E, Im T-E, Lee SJ, Jo E, Kim J, Kim SH, Chi YM, Kim J-H, Park H. 2021. The complete mitochondrial genome of Trematomus loenherbergi (Perciformes, Nototheniidae). Mitochondrial DNA B. 6:1032–1033.

Eastman JT. 1993. Antarctic fish biology: evolution in a unique environment. San Diego (CA): Academic Press.

Eastman JT. 2005. The nature of the diversity of Antarctic fishes. Polar Biol. 28:93–107.

Formenti G, Rhee A, Balacco J, Haase B, Mountcastle J, Pedrigo O, Brown S, Capodiferro MR, Al-Ajli FO, Ambrosini R, et al. 2021. Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. Genome Biol. 22:120.

Hahn G, Bachmann I, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Res. 41:e129.

Jamandre BW, Durand J-D, Tzeng W-N. 2014. High sequence variations in mitochondrial DNA control region among worldwide populations of flathead mullet Mugil cephalus. Int J Zool. 2014;564105. doi:10.1155/2014/564105.

Jemt E, Persson O, Shi Y, Mehmmedovic M, Uhler JP, López MD, Freyer C, Gustafsson CM, Samuelsson T, Falkenberg M. 2015. Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. Nucleic Acids Res. 43:9262–9275.

Lin C-Y, Lin W-W, Kao H-W. 2012. The complete mitochondrial genome of the mackerel icefish, Champsocephalus gunnari (Actinopterygii: Channichthyidae), with reference to the evolution of mitochondrial genomes in Antarctic notothenioids. Zool J Linn Soc Lond. 165:521–533.

Liu Y, Yang M, Zhou T, Xing H, Chen L, Zhang D. 2016. Complete mitochondrial genome of the Antarctic cod icefish, Pagothenna borchgrevinki (Perciformes: Nototheniidae). Mitochondrial DNA B. 1:432–433.

Papetti C, Lio P, Ruber L, Patterson J, Zardoya R. 2007. Antarctic fish mitochondrial genomes lack ND6 gene. J Mol Evol. 65:519–528.

Papetti C, Babbucci M, Dettai A, Basso A, Lucassen M, Harms L, Bonillo C, Heindler FM, Patarnello T, Negrisolo E. 2021. Not frozen in the ice: large and dynamic rearrangements in the mitochondrial genomes of the Antarctic fish. Genome Biol Evol. 13:evab017.

Ratnasingham S, Hebert PDN. 2007. BOLD: the barcode of life data system. Mol Ecol Notes. 7:355–364.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press.

Sato S, Miyabuchi M, Nishida M. 2016. Structure and variation of the mitochondrial genome of fishes. BMC Genomics. 17:719.

Sbisà E, Tazzariello F, Reyes A, Pesole G, Saccone C. 1997. Mammalian mitochondrial D-loop region structural analysis: identification of novel conserved sequences and their functional and evolutionary implications. Gene. 205:125–140.

Song W, Li L, Huang H, Zhao M, Jiang K, Zhang F, Zhao M, Chen X, Ma L. 2016. The complete mitochondrial genome sequence and gene organization of Trematomus bernacchii (Perciformes: Nototheniidae). Mol Genet Genomics. 285:457–465.

Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. In: Wolstenholme DR, editors. Mitochondrial genomes. San Diego (CA): Academic Press. p. 173–216.

Zhao X, Cheng C-HC. 2010. ND6 gene “lost” and found: evolution of mitochondrial gene rearrangements in Antarctic notothenioids. Mol Biol Evol. 27:1391–1403.