MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 4, 532–535
https://doi.org/10.1080/23802359.2024.2342932

Taylor & Francis
Taylor & Francis Group

ARTICLE COMMENTARY

Complete mitochondrial genome MK992912 of Great Knot (*Calidris tenuirostris*) is a chimera with DNA from Pacific Golden Plover *Pluvialis fulva* (Aves: Charadriiformes)

George Sangster[a] and Jolanda A. Luksenburg[b,c]

[a]Naturalis Biodiversity Center, Leiden, the Netherlands; [b]Institute of Environmental Sciences, Leiden University, Leiden, the Netherlands; [c]Department of Environmental Science and Policy, George Mason University, Fairfax, VA, USA

ABSTRACT

A complete mitochondrial genome of Great Knot (*Calidris tenuirostris*), MK992912, was published by He and colleagues in 2020. Here we show that this mitogenome is actually a chimera containing DNA fragments of both *C. tenuirostris* (15,567 bp, 92.8%) and Pacific Golden Plover (*Pluvialis fulva*, 1208 bp, 7.2%). Detecting such errors is possible before publication if each sequenced fragment is separately analyzed phylogenetically before assembling the fragments into a single mitogenome. This mitogenome has been re-used in at least four phylogenies. The error is documented to avoid the perpetuation of erroneous sequence information in the literature.

ARTICLE HISTORY
Received 6 August 2023
Accepted 9 April 2024

KEYWORDS
Chimerism; laboratory errors; mitogenome; sequence artifacts; shorebirds

Introduction

Great Knot *Calidris tenuirostris* (Horsfield, 1821) is an endangered shorebird (Charadriiformes) breeding in the tundra of northeastern Asia and wintering on coasts in southern Asia south to Australia. The first published mitochondrial genome sequence (hereafter mitogenome) of this species was published by He et al. (2020). This sequence was derived from a feather and blood sample taken from an individual caught in Rudong, Jiangsu Province, China (GenBank accession number MK992912). He et al. (2020) included a phylogram based on complete mitogenomes which placed the *C. tenuirostris* sequence sister to a sequence of Spoon-billed Sandpiper *Eurynorhynchus pygmeus*. Using a strategy described by Norén and Kullander (2018), we show that accession MK992912 is actually a chimera containing DNA of two species of shorebirds.

Materials and methods

We verified the identity of MK992912 by performing separate phylogenetic analyses of each of the two ribosomal RNA markers and the 13 protein-coding genes and comparing the position of each species in the gene trees: 12S ribosomal RNA (*12S*, 971 bp), 16S ribosomal RNA (*16S*, 1597 bp), NADH dehydrogenase subunit 1 (*ND1*, 969 bp), NADH dehydrogenase subunit 2 (*ND2*, 1041 bp), cytochrome oxidase subunit 1 (*CO1*, 1533 bp), cytochrome oxidase subunit 2 (*CO2*, 675 bp), ATP synthase subunit 8 and 6 (*ATP8-6*, 839 bp), cytochrome oxidase subunit 3 (*CO3*, 783 bp), NADH dehydrogenase subunit 3 (*ND3*, 349 bp), NADH dehydrogenase subunit 4 and 4 L (*ND4*, 1658 bp), NADH dehydrogenase subunit 5 (*ND5*, 1803 bp), cytochrome b (*CYB*, 1143 bp), and NADH dehydrogenase subunit 6 (*ND6*, 516 bp). The MITOS web server (Bernt et al. 2013) was used to obtain information on the first and last positions of individual genes. MUSCLE (as implemented in MEGA7, Kumar et al. 2016) was used to align sequences.

When we noticed that one of these data sets (*ND6*) showed a different, but strongly supported, position of *C. tenuirostris* than the other data sets, we visually compared the *C. tenuirostris* sequence with those of two other *C. tenuirostris* sequences (MK341548, Kim et al. 2020; MW160419, Chen et al. 2022) and determined the first and last positions of the anomalous fragment (Sangster and Luksenburg 2021b). We constructed separate phylogenies of (i) the anomalous fragment (1208 bp) and (ii) the rest of the mitogenome. Maximum Likelihood phylogenies were obtained using IQ-tree version 2.2.2.6 (Minh et al. 2020). We partitioned the data sets in multiple categories: ribosomal RNA, transfer RNA, protein-coding genes (each partitioned by codon), and control region. The appropriate substitution model for each partition (or appropriate combinations of partitions) in the two data sets was selected using ModelFinder (Kalyaanamoorthy et al. 2017). Branch support was obtained using ultra-fast bootstrapping (Hoang et al. 2018). Sequence divergence was calculated as uncorrected p-values with complete deletion of nucleotide positions with missing data.

The data set included the following sequences: *Alca torda* CM018102 (unpublished), *Arenaria interpres* AY074885 (Paton

CONTACT George Sangster g.sangster@planet.nl Naturalis Biodiversity Center, Darwinweg 2, PO Box 9517, 2300 RA Leiden, the Netherlands.

et al. 2002), *Calidris alba* MW168384 (Chen et al. 2022), *Calidris alpina* MW168383 (Chen et al. 2022), *Calidris canutus* MT183697(Gherardi-Fuentes et al. 2020), *Calidris pugnax* MN956840 (Chen et al. 2020), *Calidris pygmeus* KY434065 (Joen et al. 2017), *Calidris ruficollis* MG736926 (Chen et al. 2019), *Calidris subminuta* MW168385 (Chen et al. 2022), *C. tenuirostris* MK341548 (Kim et al. 2020c), *C. tenuirostris* MK992912 (He et al. 2020), *C. tenuirostris* MW160419 (Chen et al. 2022), *Charadrius alexandrinus* MF565382 (Chen et al. 2018), *Chroicocephalus brunnicephalus* JX155863 (Yang et al. 2012), *Grus grus* FJ769849 (Krajewski et al. 2010), *Pluvialis apricaria* MN122928 (unpublished), *Pluvialis fulva* KX639757 (Ding et al. 2016), *Pluvialis squatarola* MT561267 (Ding et al. 2020), *Scolopax rusticola* KM434134 (Yu et al. 2016), *Tringa erythropus* KX230491 (Cheng et al. 2016), and *Vanellus vanellus* KM577158 (Hu et al. 2016). We used *Anas poecilorhyncha* KF156760 (Zhou et al. 2015) as outgroup.

Results

Initial analysis, based on gene trees of each mitochondrial gene, showed that in the *ND6* gene tree, MK992912 clustered with Pacific Golden Plover *P. fulva*, with strong support, whereas the other gene trees placed MK992912 with two other sequences of *C. tenuirostris*. Direct (visual) comparison of the mitogenome sequences showed that the anomalous part consisted of a single 1208 bp fragment, located at positions 14,903–16,110. This represented 7.2% of the total length of MK992912 (16,775 bp) and included *ND6*. A Maximum Likelihood (ML) phylogeny of this portion of the mitogenome is shown in Figure 1(a), which shows a sister-relationship of the mitogenomes MK992912 and *P. fulva* with 100% bootstrap support. Sequence divergence between this portion of the mitogenomes MK992912 and *P. fulva* was small (1.8%). A ML phylogeny of the other parts of the mitogenome is shown in Figure 1(b), which placed MK992912 with two other sequences of *C. tenuirostris* with 100% bootstrap support.

Discussion

Our results show that different parts of MK992912 cluster with different species, each with strong bootstrap support. One of the fragments was clearly that of *P. fulva*, a similar-sized species of shorebird, which is also found in eastern Asia. A large portion of mitogenome MK992912 was nearly identical to that of two other mitogenomes of *C. tenuirostris*. MK992912 thus represents a chimera of *C. tenuirostris* and *P. fulva*.

MK992912 was obtained with Sanger sequencing. The chimera likely occurred in the laboratory resulting from the transfer of a sample of *P. fulva* to a tube intended for *C. tenuirostris* before PCR amplification or before DNA sequencing. Indeed, a mitogenome of *P. fulva* was sequenced by members of the same team and was published several years previously (Ding et al. 2016). Detecting such errors is possible if each fragment is separately analyzed phylogenetically before assembling the fragments into a single mitogenome.

We report this problematic mitogenome because accumulation of erroneous sequences may compromise subsequent applications, including DNA identification, primer design for intraspecific studies, phylogenetic inference, historical biogeography, taxonomy, and comparative analysis (Sangster and Luksenburg 2021c). Indeed, we found four re-uses of the mitogenome (Ding et al. 2020; Gherardi-Fuentes et al. 2020; Guo et al. 2021; Yang et al. 2021). In all four cases, the mitogenome was included in a phylogeny. Černý and Natale (2022) used several non-problematic fragments of MK992912 but removed the problematic ND6 fragment from their analyses.

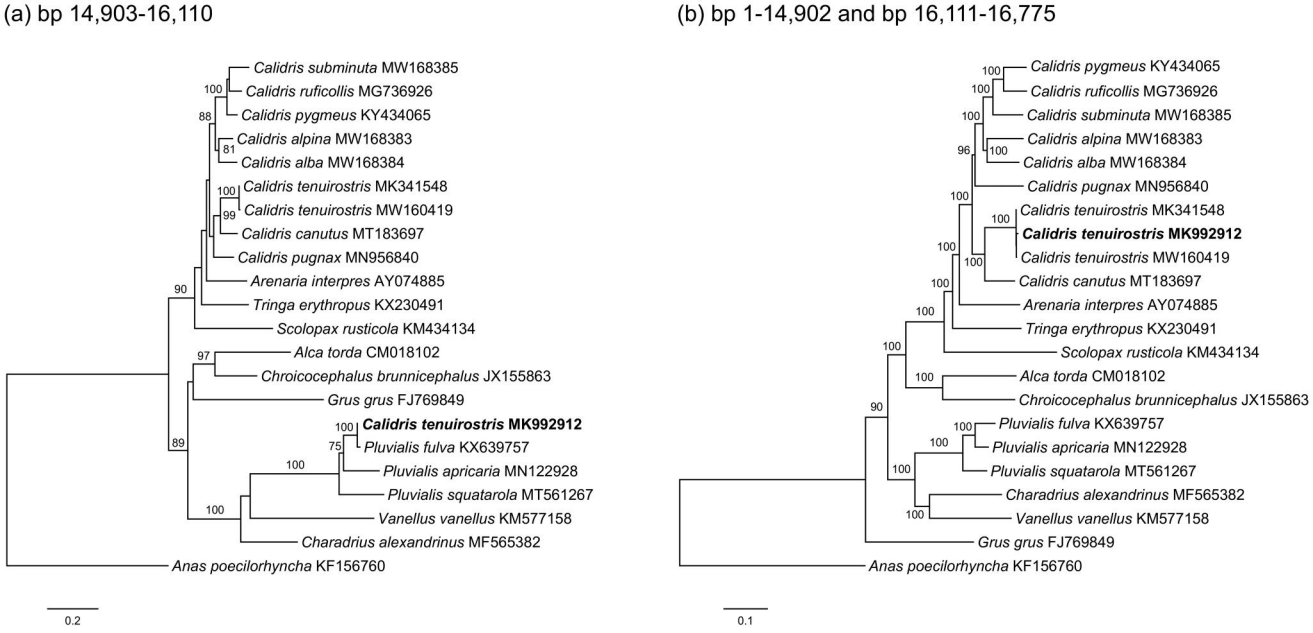

Figure 1. ML phylogenies of shorebirds (Charadriiformes) based on (a) positions 14,903–16,110 (1208 bp) of the mitogenome, (b) mitogenomes excluding positions 14,903–16,110. Numbers along branches represent bootstrap support values (>70%) based on 1000 pseudoreplications. Note the different position of *C. tenuirostris* (MK992912) in the two gene trees.

Our study, and that of Sangster and Luksenburg (2023), underscore that chimeras are easily overlooked without dedicated analysis of the entire mitogenomes. We suspect that the few cases of chimerism reported so far in vertebrate mitogenomics (e.g. Norén and Kullander 2018; Sangster and Luksenburg 2020, 2021a, 2021b, 2023) do not reflect the true prevalence of this problem. Clearly, greater vigilance is necessary during laboratory procedures, quality control of raw data, and peer review of the final sequences.

Acknowledgment

The authors are grateful for the constructive comments of the four reviewers and the Associate Editor.

Author contributions

George Sangster: conceptualization, methodology, formal analysis, investigation, and writing—original draft. Jolanda Luksenburg: writing—review and editing. Both authors agreed to be accountable for all aspects of the work and approved the final draft to be published.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was received for this study.

ORCID

George Sangster http://orcid.org/0000-0002-2475-7468
Jolanda A. Luksenburg http://orcid.org/0000-0003-4424-4368

Data availability statement

No new sequence data were generated for this study.

References

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69(2):313–319. doi:10.1016/j.ympev.2012.08.023.

Černý D, Natale R. 2022. Comprehensive taxon sampling and vetted fossils help clarify the time tree of shorebirds (Aves, Charadriiformes). Mol Phylogenet Evol. 177:107620. doi:10.1016/j.ympev.2022.107620.

Chen TY, Zhang F, Wang G-H, Zhang M-Y, Liang T, Lu CH. 2020. Complete mitochondrial genome of the Ruff, *Calidris pugnax* (Aves, Scolopacidae). Mitochondrial DNA B Resourg. 5(2):1246–1247. doi:10.1080/23802359.2020.1731356.

Chen W, Liu W, Zhang C, Li K, Hu C, Chang Q. 2019. The complete mitochondrial genome of the Red-necked Stint *Calidris ruficollis* (Charadriiformes, Scolopacidae). Conservation Genet Resour. 11(2):181–184. doi:10.1007/s12686-018-0996-1.

Chen W, Miao K, Wang J, Wang H, Sun W, Yuan S, Luo S, Hu C, Chang Q. 2022. Five new mitogenomes sequences of Calidridine sandpipers (Aves: Charadriiformes) and comparative mitogenomics of genus *Calidris*. PeerJ. 10:e13268. doi:10.7717/peerj.13268.

Chen W, Zhang C, Pan T, Liu W, Li K, Hu C, Chang Q. 2018. The mitochondrial genome of the Kentish Plover *Charadrius alexandrinus* (Charadriiformes: Charadriidae) and phylogenetic analysis of Charadrii. Genes Genomics. 40(9):955–963. doi:10.1007/s13258-018-0703-3.

Cheng Y, Zhou L, Dong Y. 2016. Complete mitochondrial genome of *Tringa erythropus* (Charadriiformes: Scolopacidae). Mitochondrial DNA B Resour. 1(1):678–679. doi:10.1080/23802359.2016.1219637.

Ding J, Liu W, Zhang Y, Chang Q, Hu C. 2016. The complete mitochondrial genome of Pacific Golden Plover *Pluvialis fulva* (Charadriiformes, Charadriidae). Mitochondrial DNA B Resour. 1(1):701–702. doi:10.1080/23802359.2016.1225524.

Ding J, Qian R, Tai D, Yao W, Hu C, Chang Q. 2020. The complete mitochondrial genome of Grey Plover *Pluvialis squatarola* (Charadriiformes, Charadriidae). Mitochondrial DNA B Resour. 5(3):2738–2739. doi:10.1080/23802359.2020.1787892.

Gherardi-Fuentes C, Navedo JG, Verdugo C. 2020. The mitochondrial genome of two long-distance migratory shorebirds: the Hudsonian Godwit (*Limosa haemastica*) and the Red knot (*Calidris canutus*). Mitochondrial DNA B Resour. 5(3):3553–3554. doi:10.1080/23802359.2020.1827997.

Guo T, Yang Z, Zhu Y, Wang H, Kong D, Hu C, Liu M. 2021. The complete mitochondrial genome of Eastern Curlew *Numenius madagascariensis* (Charadriiformes, Scolopacidae). Mitochondrial DNA B Resour. 6(11):3091–3092. doi:10.1080/23802359.2021.1959448.

He Y, Liu W, Wu Y, Chang Q. 2020. The complete mitochondrial genome of *Calidris tenuirostris* (Charadriiformes: Scolopacidae). Mitochondrial DNA B Resour. 5(1):178–179. doi:10.1080/23802359.2019.1698375.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 35(2):518–522. doi:10.1093/molbev/msx281.

Hu C, Yu T, Xie W, Zhu Y, Chang Q. 2016. The complete mitochondrial genome of *Vanellus vanellus* (Charadriiformes: Charadriidae). Mitochondrial DNA A DNA Mapp Seq Anal. 27(3):1936–1937. doi:10.3109/19401736.2014.971295.

Joen HS, Lee MY, Choi YS, An J. 2017. Mitochondrial genome analysis of the spoon-billed sandpiper (*Eurynorhynchus pygmeus*). Mitochondrial DNA B Resour. 2(1):150–151. doi:10.1080/23802359.2017.1298415.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14(6):587–589. doi:10.1038/nmeth.4285.

Kim SI, Lee MY, Jeon HS, Cha I, Park HO, Yeo KW, An J. 2020. Complete mitochondrial genome of two shorebirds (Charadriiformes: Scolopacidae), Great Knot (*Calidris tenuirostris*) and Bar-tailed Godwit (*Limosa lapponica*). Mitochondrial DNA B Resour. 5(3):3516–3518. doi:10.1080/23802359.2020.1750976.

Krajewski C, Sipiorski JT, Anderson FE. 2010. Complete mitochondrial genome sequences and the phylogeny of cranes (Gruiformes: Gruidae). Auk. 127(2):440–452. doi:10.1525/auk.2009.09045.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 33(7):1870–1874. doi:10.1093/molbev/msw054.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37(5):1530–1534. doi:10.1093/molbev/msaa015.

Norén M, Kullander S. 2018. The enigmatic *Betadevario ramachandrani* (Teleostei: Cyprinidae: Danioninae): phylogenetic position resolved by mitogenome analysis, with remarks on the prevalence of chimeric mitogenomes in GenBank. Cogent Biol. 4(1):1525857. doi:10.1080/23312025.2018.1525857.

Paton TA, Haddrath O, Baker AJ. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc Biol Sci. 269(1493):839–846. doi:10.1098/rspb.2002.1961.

Sangster G, Luksenburg JA. 2020. The published complete mitochondrial genome of *Eptesicus serotinus* is a chimera of *Vespertilio sinensis* and *Hypsugo alaschanicus* (Mammalia: Chiroptera). Mitochondrial DNA B Resour. 5(3):2661–2664. doi:10.1080/23802359.2020.1785349.

Sangster G, Luksenburg JA. 2021a. Scientific data laundering: chimeric mitogenomes of a sparrowhawk and a nightjar covered-up by forged phylogenies. Biochem Syst Ecol. 96:104263. doi:10.1016/j.bse.2021.104263.

Sangster G, Luksenburg JA. 2021b. Sharp increase of problematic mitogenomes of birds: causes, effects and remedies. Gen Biol Evol. 13(9): evab210. doi:10.1093/gbe/evab210.

Sangster G, Luksenburg JA. 2021c. The published complete mitochondrial genome of milk shark (*Rhizoprionodon acutus*) is a misidentified Pacific spadenose shark (*Scoliodon macrorhynchos*) (Chondrichthyes: Carcharhiniformes). Mitochondrial DNA B Resour. 6(3):828–830. doi:10.1080/23802359.2021.1884019.

Sangster G, Luksenburg JA. 2023. The published complete mitochondrial genome of Spotted Greenshank (*Tringa guttifer*) is a chimera with DNA from Red-necked Stint (*Calidris ruficollis*) (Aves: Charadriiformes). Mitochondrial DNA B Resour. 8(11):1273–1275. doi:10.1080/23802359.2023.2282791.

Yang C, Hou X, Liu BY, Gong HS, Yuan H, Li XJ, Tang J, Wang Y. 2021. The mitogenome of Common Snipe, *Gallinago Gallina go gallinago* Linnaeus, 1758 and evolutionary implications for the family Scolopacidae. Mitochondrial DNA B Resour. 6(10):2886–2889. doi:10.1080/23802359.2021.1972870.

Yang C, Wang QX, Huang Y, Xiao H. 2012. Analysis of the complete mitochondrial genome sequence of *Larus brunnicephalus* (Aves, Laridae). Yi Chuan. 34(11):1434–1446. doi:10.3724/sp.j.1005.2012.01434.

Yu T, Hu C, Chang Q, Yang R, Dong W, Xie W. 2016. The complete mitochondrial genome of *Scolopax rusticola* (Charadriiformes: Scolopacidae). Mitochondrial DNA A DNA Mapp Seq Anal. 27(3):1800–1801. doi:10.3109/19401736.2014.963819.

Zhou W, Zhang C, Pan T, Yan L, Hu C, Xue C, Chang Q, Zhang B. 2015. The complete mitochondrial genome of *Anas poecilorhyncha* (Anatidae: Anas). Mitochondrial DNA. 26(2):265–266. doi:10.3109/19401736.2013.823191.