

Comment

Contamination of Consensus Sequences in Next-Generation Mitogenomics: The Published Mitochondrial Genome of *Haliastur indus* Is a Chimera with DNA from *Butastur indicus*. Comment on Sonongbua et al. Insights into Mitochondrial Rearrangements and Selection in Accipitrid Mitogenomes, with New Data on *Haliastur indus* and *Accipiter badius poliopsis*. *Genes* 2024, 15, 1439

George Sangster^{1,*}  and Jolanda A. Luksenburg^{2,3}¹ Naturalis Biodiversity Center, Darwinweg 2, 2300 RA Leiden, The Netherlands² Institute of Environmental Sciences, Leiden University, Einsteinweg 2, 2333 CC Leiden, The Netherlands³ Department of Environmental Science and Policy, George Mason University, 4400 University Drive, Fairfax, VA 22030 4444, USA

* Correspondence: g.sangster@planet.nl

The quality and authenticity of DNA sequences have long been points of concern for molecular biologists and systematists [1–3]. As the field of mitogenomics moves from traditional (Sanger) sequencing to next-generation sequencing (NGS), new challenges may arise and old problems may present themselves in new ways. While multiple chimeric mitogenomes that were generated using Sanger sequencing have been documented [4–11], so far no chimeras appear to have been identified that were generated using NGS. Here we provide evidence that a recently published mitochondrial genome that was generated using NGS is an interspecific chimera, demonstrating that this is likely an overlooked problem in next-generation mitogenomics.

A complete mitochondrial genome of the Brahminy Kite *Haliastur indus* (Aves: Accipitridae) was published as GenBank accession number OP133375.1/NC_066800.1 and documented, along with a mitochondrial genome of *Accipiter badius*, in a detailed paper in *Genes* [12]. The DNA sequence was generated using an Illumina Novaseq™ 6000 from a blood sample of a live bird obtained in Samut Prakarn, Thailand. This is the first mitogenome of *H. indus* and thus has the potential to become the reference sequence for future mitochondrial DNA applications.

We assessed the phylogenetic placement of OP133375.1 among Accipitriformes using three mitochondrial markers that are commonly used in avian phylogenetics: NADH dehydrogenase subunit 2 (*ND2*, 1041 bp), part of cytochrome oxidase subunit 1 (*CO1*, 696 bp), and cytochrome b (*CYB*, 1143 bp). The MITOS2 web server [13] was used to obtain information on the first and last positions of individual genes. MUSCLE [14] (as implemented in MEGA7 [15]) was used to align sequences. Maximum Likelihood phylogenies were obtained using IQ-tree version 2.2.2.6 [16]. We partitioned the data sets by codon. The appropriate substitution model for each partition in the data sets was selected using ModelFinder [17]. Branch support was obtained using 1000 ultra-fast bootstraps [18]. We used a sequence of *Cathartes melambrotus* (PQ153910, [19]) as an outgroup based on previous studies of Accipitriformes [20,21]. Sequence divergence was



Academic Editors: Selvarangan Ponnazhagan and Christian Roos

Received: 8 April 2025

Revised: 16 December 2025

Accepted: 22 December 2025

Published: 14 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

calculated as uncorrected p -values with complete deletion of nucleotide positions with missing data.

Following [6], we looked for evidence of chimeras based on one or more of five possible signs: (i) placement on a long branch, (ii) a deep divergence, (iii) a distant position from its expected phylogenetic position, (iv) a sister species placed on a long branch, (v) a mismatch between gene trees. An anomalous sequence was visually inspected and compared with sequences of conspecifics to localize any divergent fragment(s).

In the *ND2* gene tree, OP133375.1 was sister to a reference sequence of *H. indus* with 100% bootstrap support but the two sequences were rather divergent (as shown by their relatively long branches; Figure 1a). Uncorrected sequence divergence between the two sequences was 3.7%. In the *CO1* gene tree, OP133375 clustered with two reference sequences of *H. indus* with a very small sequence divergence (0.0–0.3%; Figure 1b). In the *CYB* gene tree, OP133375.1 clustered with three reference sequences of *Butastur indicus* with high bootstrap support (100%; Figure 1c). The *CYB* fragment of OP133375.1 was identical to that of one of the *B. indicus* sequences (AB830616), and differed from two other reference sequences by 0.5–0.6%. Sequence divergence between OP133375.1 and a single *CYB* reference sequence of *H. indus* (AY987309) was 10.9%.

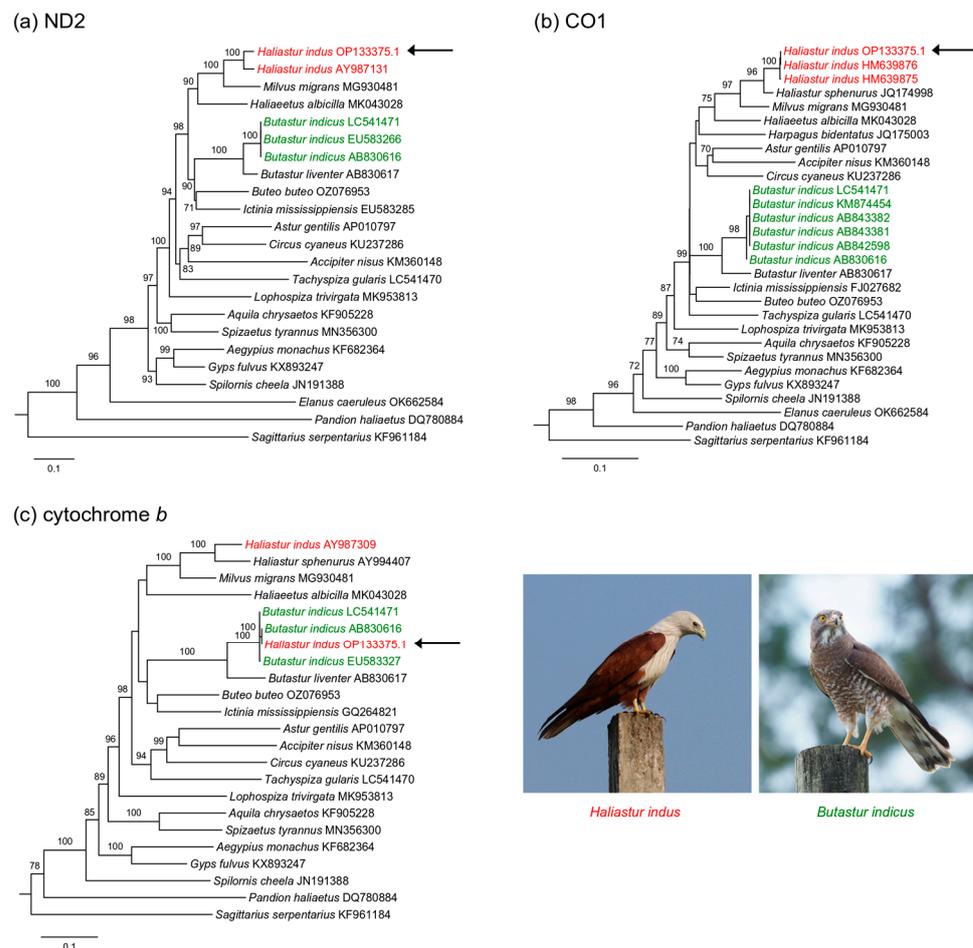


Figure 1. Maximum Likelihood phylogenies of Accipitriformes based on (a) *ND2* (1041 bp), (b) *CO1* (696 bp), and (c) *CYB* (1143 bp) sequences. The outgroup (*Cathartes melambrotus*) is not shown. Numbers along branches represent bootstrap support values (>70%) based on 1000 pseudoreplications. The arrows indicate the problematic sequence of *Haliastur indus*. Photograph of Brahminy Kite (*Haliastur indus*) by Shino Jacob Kootanad (CC-BY-SA 4.0, Wikimedia); photograph of Grey-faced Buzzard (*Butastur indicus*) by nhm6306 (CC-BY-NC, iNaturalist).

A large fragment of OP133375.1 was identical to one of the *B. indicus* sequences (AB830616) and differed 1.3% from another mitogenome of *B. indicus* (MW017133). This fragment comprised 8994 bp and spanned the positions 1–1690 and 11,752–19,055 of OP133375; this corresponded to 47.2% of the total length of the mitogenome.

Our data show that OP133375.1 was present at two different well-supported phylogenetic positions in mitochondrial gene trees. This is a clear indication that this sequence is a chimera comprising DNA from at least two different species. The *CO1* gene tree indicates that part of the mitogenome is referable to *H. indus* and the *CYB* gene tree shows that the mitogenome also includes DNA fragments of *B. indicus*. Our data further show that a long fragment of mitogenome OP133375.1 is identical to a mitogenome sequence of *B. indicus*. Consequently, mitogenome OP133375.1 is not an accurate representation of the true mitogenome of *H. indus* and should not be used in biological studies.

Given that the mitogenome was obtained with NGS technology [12], the only plausible explanation for the observed sequence abnormalities is that these are due to failure to remove a reference sequence (almost certainly *B. indicus* AB830616) during the preparation of the consensus sequence.

Our study shows, perhaps for the first time, that chimeras do occur in mitogenomes assembled from NGS-generated data. Whereas in traditional (Sanger) sequencing work chimeras result from errors during the ‘wet’ phase (i.e., in the laboratory due to the inadvertent transfer of a DNA extract or PCR product to the wrong tube), NGS chimeras may be generated in the ‘dry’ phase (i.e., while generating a consensus sequence during the bioinformatics phase of the study).

The cladogram in [12] showed *H. indus* in its expected phylogenetic position, i.e., sister to Black Kite *Milvus migrans* [20,21]. This is quite remarkable given that almost 50% of its sequence is derived from a more distantly related species. Thus, a chimera may be difficult to detect in a phylogeny that includes no other conspecific sequences. We were able to detect this chimera because we generated separate gene trees of three markers, and because the problematic parts of the mitogenome referred to another, not particularly closely related species, and overlapped with one of the three reference markers. The detection of chimeras will be almost impossible if (i) a reference sequence of a conspecific individual is used, or (ii) the heterospecific fragment(s) do not overlap with at least one of the reference markers.

Large numbers of mitogenomes are now routinely generated for a single study using NGS technology [22–27]. The generation of the consensus sequence may be regarded as the Achilles heel of NGS mitogenomics because a seemingly small error (failure to remove the reference sequence) may have far-reaching consequences. Awareness of the problem of chimeras is important because chimeras may compromise research studies, including studies that re-use one or more mitogenomes for other purposes. For instance, previously published chimeric sequences have led to proposals for unwarranted taxonomic revision and may have compromised comparative analyses and measurements of divergence times [6]. We hope that our commentary will draw attention to the risk of generating chimeric sequences and will help others detect and document additional cases of this likely overlooked problem.

Author Contributions: Conceptualization, writing—draft preparation: G.S., Review and editing: J.A.L. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Clark, A.G.; Whittam, T.S. Sequencing errors and molecular evolutionary analysis. *Mol. Biol. Evol.* **1992**, *9*, 744–752. [[CrossRef](#)] [[PubMed](#)]

2. Arctander, P. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proc. Biol. Sci.* **1995**, *262*, 13–19. [[PubMed](#)]
3. Helbig, A.J.; Seibold, I. Are storks and new world vultures paraphyletic? *Mol. Phylogenet. Evol.* **1996**, *6*, 315–319. [[CrossRef](#)] [[PubMed](#)]
4. Norén, M.; Kullander, S. 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.* **2018**, *4*, 1525857. [[CrossRef](#)]
5. Sangster, G.; Luksenburg, J.A. The published complete mitochondrial genome of *Eptesicus serotinus* is a chimera of *Vespertilio sinensis* and *Hypsugo alaschanicus* (Mammalia: Chiroptera). *Mitochondrial DNA Part B* **2020**, *5*, 2661–2664. [[CrossRef](#)]
6. Sangster, G.; Luksenburg, J.A. Sharp increase of problematic mitogenomes of birds: Causes, effects and remedies. *Gen. Biol. Evol.* **2021**, *13*, evab210. [[CrossRef](#)]
7. Sangster, G.; Luksenburg, J.A. Scientific data laundering: Chimeric mitogenomes of a sparrowhawk and a nightjar covered-up by forged phylogenies. *Biochem. Syst. Ecol.* **2021**, *96*, 104263. [[CrossRef](#)]
8. Sangster, G.; Luksenburg, J.A. 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 Part B* **2023**, *8*, 1273–1275. [[CrossRef](#)]
9. Sangster, G.; Luksenburg, J.A. Complete mitochondrial genome MK992912 of Great Knot (*Calidris tenuirostris*) is a chimera with DNA from Pacific Golden Plover *Pluvialis fulva* (Aves: Charadriiformes). *Mitochondrial DNA Part B* **2024**, *9*, 532–535. [[CrossRef](#)]
10. Sangster, G.; Luksenburg, J.A. The published complete mitochondrial genome of Blue-fronted Redstart (*Phoenicurus frontalis*) is a chimera and includes DNA from Pink-rumped Rosefinch *Carpodacus waltoni eos* (Aves: Passeriformes). *Mitochondrial DNA Part B* **2024**, *9*, 861–864. [[CrossRef](#)]
11. Islam, S.; Peart, C.; Kehlmaier, C.; Sun, Y.H.; Lei, F.; Dahl, A.; Klemroth, S.; Alexopoulou, D.; Delgado, M.d.M.; Laiolo, P.; et al. Museomics help resolving the phylogeny of snowfinches (Aves, Passeridae, *Montifringilla* and allies). *Mol. Phylogenet. Evol.* **2024**, *198*, 108135. [[CrossRef](#)] [[PubMed](#)]
12. Sonongbua, J.; Thong, T.; Panthum, T.; Budi, T.; Singchat, W.; Kraichak, E.; Chaiyes, A.; Muangmai, N.; Duengkae, P.; Sitdhibutr, R.; et al. Insights into mitochondrial rearrangements and selection in accipitrid mitogenomes, with new data on *Haliastur indus* and *Accipiter badius poliopsis*. *Genes* **2024**, *15*, 1439. [[CrossRef](#)] [[PubMed](#)]
13. Bernt, M.; Donath, A.; Jühling, F.; Externbrink, F.; Florentz, C.; Fritzsche, G.; Pütz, J.; Middendorf, M.; Stadler, P.F. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **2013**, *69*, 313–319. [[CrossRef](#)]
14. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucl. Acid. Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
15. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
16. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; Von Haeseler, A.; Lanfear, R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)]
17. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermini, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [[CrossRef](#)]
18. Hoang, D.T.; Chernomor, O.; von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522. [[CrossRef](#)]
19. Driver, R.; Li, R. The complete mitochondrial genome and phylogenetic analyses of *Cathartes melambrotus* (Wetmore 1964) (Aves: Cathartidae). *Mitochondrial DNA Part B* **2025**, *10*, 187–191. [[CrossRef](#)]
20. Mindell, D.P.; Fuchs, J.; Johnson, J.A. Phylogeny, taxonomy, and geographic diversity of diurnal raptors: Falconiformes, Accipitriformes, and Cathartiformes. In *Birds of Prey*; Sarasola, J.H., Grande, J., Negro, J., Eds.; Springer: Cham, Switzerland, 2018; pp. 3–32.
21. Catanach, T.A.; Halley, M.R.; Pirro, S. Enigmas no longer: Using ultraconserved elements to place several unusual hawk taxa and address the non-monophyly of the genus *Accipiter* (Accipitriformes: Accipitridae). *Biol. J. Linn. Soc.* **2025**, *144*, blae028. [[CrossRef](#)]
22. Stryjewski, K.F.; Sorenson, M. Mosaic genome evolution in a recent and rapid avian radiation. *Nat. Ecol. Evol.* **2017**, *1*, 1912. [[CrossRef](#)] [[PubMed](#)]
23. Krzemińska, U.; Morales, H.E.; Greening, C.; Nyári, Á.S.; Wilson, R.; Song, B.K.; Austin, C.M.; Sunnucks, P.; Pavlova, A.; Rahman, S. Population mitogenomics provides insights into evolutionary history, source of invasions and diversifying selection in the House Crow (*Corvus splendens*). *Heredity* **2018**, *120*, 296–309. [[CrossRef](#)]
24. Wang, E.; Zhang, D.; Braun, M.S.; Hotz-Wagenblatt, A.; Pärt, T.; Arlt, D.; Schmaljohann, H.; Bairlein, F.; Lei, F.; Wink, M. Can mitogenomes of the Northern Wheatear (*Oenanthe oenanthe*) reconstruct its phylogeography and reveal the origin of migrant birds? *Sci. Rep.* **2020**, *10*, 9290. [[CrossRef](#)]

25. Lombardo, G.; Rambaldi Migliore, N.; Colombo, G.; Capodiferro, M.R.; Formenti, G.; Caprioli, M.; Moroni, E.; Caporali, L.; Lancioni, H.; Secomandi, S.; et al. The mitogenome relationships and phylogeography of Barn Swallows (*Hirundo rustica*). *Mol. Biol. Evol.* **2022**, *39*, msac113. [[CrossRef](#)]
26. Al-Jumaili, A.S.; Hanotte, O. The usefulness of maternally inherited genetic markers for phylogeographic studies in village chicken. *Anim. Biotechnol.* **2023**, *34*, 863–881. [[CrossRef](#)] [[PubMed](#)]
27. Cloutier, A.; Chan, D.T.C.; Poon, E.S.K.; Sin, S.Y.W. The genetic consequences of historic climate change on the contemporary population structure of a widespread temperate North American songbird. *Mol. Phylogenet. Evol.* **2024**, *201*, 108216. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.