




BRIEF REPORT

REVISED Intraspecific mitochondrial gene variation can be as low as that of nuclear rRNA [version 2; peer review: 2 approved]

Tshifhiwa G. Matumba^{1,2}, Jody Oliver¹, Nigel P. Barker³, Christopher D. McQuaid², Peter R. Teske ¹

¹Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Auckland Park, 2006, South Africa

²Department of Zoology and Entomology, Rhodes University, Grahamstown, 6140, South Africa

³Department of Plant and Soil Sciences, University of Pretoria, Hatfield, 0028, South Africa

V2 First published: 07 May 2020, 9:339
<https://doi.org/10.12688/f1000research.23635.1>
 Latest published: 28 Aug 2020, 9:339
<https://doi.org/10.12688/f1000research.23635.2>

Abstract

Background: Mitochondrial DNA (mtDNA) has long been used to date historical demographic events. The idea that it is useful for molecular dating rests on the premise that its evolution is neutral. Even though this idea has long been challenged, the evidence against clock-like evolution of mtDNA is often ignored. Here, we present a particularly clear and simple example to illustrate the implications of violations of the assumption of selective neutrality.





Methods: DNA sequences were generated for the mtDNA COI gene and the nuclear 28S rRNA of two closely related rocky shore snails, and species-level variation was compared. Nuclear rRNA is not usually used to study intraspecific variation in species that are not spatially structured, presumably because this marker is assumed to evolve so slowly that it is more suitable for phylogenetics.

Results: Even though high inter-specific divergence reflected the faster evolutionary rate of COI, intraspecific genetic variation was similar for both markers. As a result, estimates of population expansion times based on mismatch distributions differed between the two markers by millions of years.

Conclusions: Assuming that 28S evolution is more clock-like, these findings can be explained by variation-reducing purifying selection in mtDNA at the species level, and an elevated divergence rate caused by diversifying selection between the two species. Although these two selective forces together make mtDNA suitable as a marker for species identifications by means of DNA barcoding because they create a 'barcoding gap', estimates of demographic change based on this marker can be expected to be highly unreliable. Our study contributes to the growing evidence that the utility of mtDNA sequence data beyond DNA barcoding is limited.

Open Peer Review

Reviewer Status  

	Invited Reviewers	
	1	2
version 2 (revision) 28 Aug 2020		 report
version 1 07 May 2020	 report	  report

1. **Andrew G. Briscoe**, Natural History Museum, London, UK

2. **Abigail Hui En Chan**, Mahidol University, Bangkok, Thailand

Urusa Thaenkham, Mahidol University, Bangkok, Thailand

Any reports and responses or comments on the article can be found at the end of the article.

Keywords

Purifying selection, diversifying selection, mismatch distribution, molecular dating, demographic history, population expansion

Corresponding author: Peter R. Teske (pteske101@gmail.com)

Author roles: **Matumba TG:** Formal Analysis, Investigation, Methodology; **Oliver J:** Formal Analysis; **Barker NP:** Supervision, Writing – Review & Editing; **McQuaid CD:** Conceptualization, Supervision, Writing – Review & Editing; **Teske PR:** Formal Analysis, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work is based upon research supported by the National Research Foundation of South Africa (Grant number 64801 to N.P.B) and Rhodes University.

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

Copyright: © 2020 Matumba TG *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Matumba TG, Oliver J, Barker NP *et al.* **Intraspecific mitochondrial gene variation can be as low as that of nuclear rRNA [version 2; peer review: 2 approved]** F1000Research 2020, 9:339 <https://doi.org/10.12688/f1000research.23635.2>

First published: 07 May 2020, 9:339 <https://doi.org/10.12688/f1000research.23635.1>

REVISED Amendments from Version 1

This is a minor update. Some sentences have been modified (including in the abstract) and several additional references have been included.

Any further responses from the reviewers can be found at the end of the article

Introduction

Mitochondrial DNA (mtDNA) has long been a marker of choice for investigating concepts as diverse as estimating genetic diversity and effective population sizes, reconstructing species' evolutionary histories, exploring spatial genetic subdivisions, and identifying cryptic species. All these methods assume that mtDNA variation conforms to the neutral model of molecular evolution¹, but violations of this premise have long been recognised². Over the past decades, much evidence has accumulated that mtDNA can be strongly affected by selective sweeps and background selection³⁻⁶. As a result, the usefulness of the marker in assessing genetic diversity⁷ and exploring spatial genetic structure in continuously distributed populations⁸ has been questioned, and corrections of the mitochondrial molecular clock that account for selection have been proposed^{9,10}.

The implications of reduced genetic diversity at the species or population levels due to purifying selection has so far received little attention. When mutations in mitochondrial genes occur at fewer sites than expected under the neutral model¹¹, molecular dating of historical demographic events by means of evolutionary rate estimates that are typically based on interspecific divergence^{12,13} will result in considerable underestimates. This is particularly likely because divergence between species can be strongly affected by diversifying selection that is driven by different environmental conditions^{14,15}, resulting in a faster accumulation of mutations characterising each species than is expected under the neutral model.

Here, we explore this issue using mitochondrial and nuclear DNA sequence data from two common southern African snails of the genus *Afrolittorina* that show no spatial genetic structure throughout their ranges¹⁶. The finding that data from two genetic markers with mutation rates that are assumed to differ by at least an order of magnitude^{17,18} have similar levels of intraspecific variation challenges the usefulness of mitochondrial DNA sequences for studying historical demographic changes.

Methods

Specimens of the snails *Afrolittorina africana* and *A. knysnaensis* were collected at 34 sites throughout South Africa (Table 1). DNA was extracted using the CTAB protocol¹⁹, amplified with universal COI primers²⁰ and 28S primers LSU5²¹ and LSU1600²² following Williams *et al.*²², and sequenced on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) using Big Dye Terminator v3.1 chemistry. Sequences were edited using MEGA7²³, and 28S sequences were phased in PHASE v2.1.1²⁴ using default settings. Genealogical relationships between COI haplotypes and 28S alleles were reconstructed using the median-joining algorithm²⁵ in popArt v1.7²⁶. To explore the effect of using interspecific evolutionary rates to estimate

species-level population size changes, we calculated population expansion time²⁷ using Arlequin v3.5²⁸ using each marker's slowest and fastest published rates for marine gastropods (Table 2).

Table 1. Number of individuals of *Afrolittorina africana* and *A. knysnaensis* for which COI and 28S sequences were generated. 34 sites along the South African coastline were sampled, and these are arranged from west to east.

Site name	<i>A. africana</i>	<i>A. knysnaensis</i>
Port Nolloth	-	4
Groenriviersmond	-	5
Strandfonteinpunt	-	1
Lamberts Bay	-	5
Melkbosstrand	-	2
Paternoster	-	1
Yzerfontein	-	1
Rooiels	-	2
Cape Agulhas	-	2
Still Bay	1	2
Herolds Bay	-	2
The Wilderness	2	-
Sedgefield	1	1
Tsitsikamma	-	1
Jeffreys Bay	-	3
Cape Recife	-	8
Cannon Rocks	6	4
Bushmans River	3	4
Port Alfred	7	4
Fish River	6	-
Hamburg	2	-
Gqunube	4	2
Haga-Haga	8	10
Dwesa	6	-
Hluleka	2	6
Port St Johns	5	6
Port Edward	8	2
Ramsgate	4	4
Park Rynie	2	-
Mhlanga	4	-
Ballito	3	-
Sheffield	12	-
Zinkwazi	3	-
Mission Rocks	4	-

Results

Species-specific genetic clusters reconstructed from COI sequences were highly distinct (Figure 1a), with a minimum number of 44 nucleotide differences between the two species' most closely related haplotypes. In contrast, differentiation between 28S sequences (Figure 1b) was an order of magnitude smaller (4 differences).

In contrast to the high inter-specific differentiation between COI haplotypes, intra-specific genetic differentiation was comparatively low for this marker, and similar to that of 28S. In

A. knysnaensis, six COI haplotypes and seven 28S haplotypes were found, while the maximum differentiation between the COI haplotypes was only two nucleotide differences, but 10 for 28S. The number of haplotypes was greater for *A. africana*, where 14 were found for COI and 10 for 28S. Maximum nucleotide differences for this species were seven in the COI network and five for 28S.

The practical implications of two markers with very different evolutionary rates based on inter-specific divergence having similar levels of intraspecific variation are illustrated in Table 2.

Table 2. Estimates of population expansions of the two species of *Afrolittorina* under the sudden expansion model. The moment estimator τ is equal to $2ut$, where u equals $2\mu k$ (μ is the mutation rate and k is the length of the sequence), and t is the time of expansion in million of years (my).

Species	τ	Marker	μ (%.my ⁻¹)	t (my)
<i>Afrolittorina knysnaensis</i>	2.00	COI	0.50 ¹	0.40 (0.00 – 0.41)
			2.60 ²	0.07 (0.00 – 0.08)
	3.25	28S	0.01 ¹	32.1 (18.5 – 61.3)
			0.05 ²	6.41 (3.69 – 12.3)
<i>Afrolittorina africana</i>	2.50	COI	0.50 ¹	0.50 (0.30 – 0.79)
			2.60 ²	0.10 (0.06 – 0.15)
	2.75	28S	0.01 ¹	27.1 (19.1 – 51.5)
			0.05 ²	5.42 (3.81 – 10.3)

¹Malaquias & Reid 2009¹⁸; ²Williams & Reid 2004¹⁷. A generation time of one year was used.

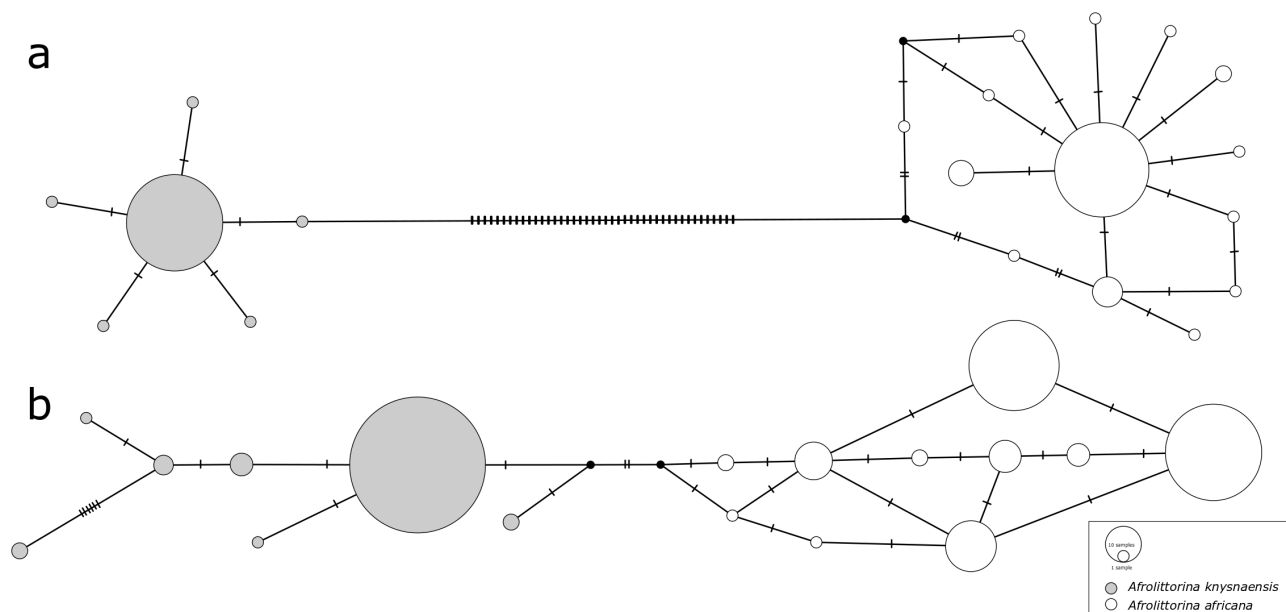


Figure 1. Median-joining haplotype networks constructed from **a**) COI sequences and **b**) 28S rRNA sequences of *Afrolittorina knysnaensis* (grey) and *A. africana* (white). Low intra-specific variation and high inter-specific variation of COI potentially illustrate purifying and diversifying selection, respectively. The size of circles is proportional to the frequency of each haplotype, cross-bars represent nucleotide differences, and black dots are missing haplotypes not found in the samples.

Using published rates, estimates of population expansion times were more than an order of magnitude greater based on the 28S data than based on the COI data.

Discussion

The usefulness of the mtDNA COI gene to uncover overlooked biodiversity is undisputed because of the marker's tendency to have a well-defined barcoding gap, as was found here. The two study species' COI sequences were much more strongly differentiated than their 28S sequences, potentially reflecting diversifying selection as a result of adaption to different thermal environments¹⁶. In contrast, there was comparatively little genetic variation at the intraspecific level for either marker, which is likely due to the commonly reported strong purifying selection acting upon the COI gene^{6,9}.

Many researchers explore their mtDNA sequence data for additional information, but the selective forces that together create the barcoding gap²⁹ make its utility for other applications questionable^{7,8}. In the present study, we have highlighted a largely unexplored problem that likely arises from selection effects in mtDNA data: the fact that demographic events using gene regions under variation-reducing purifying selection are dated using molecular clock calibrations affected by variation-increasing diversifying selection. The finding that intraspecific mtDNA variation can be as low as that of nuclear rRNA cautions against the continued use of mtDNA for exploring demographic trends

by means of mismatch distributions or Bayesian skyline plots³⁰, a practice that continues to dominate the recent literature^{31–34}.

In our opinion, it is time to discontinue the use of fixed mtDNA rates based on divergence dating of closely related taxa, such as the closure of the Central American Seaway to date phylogenies of marine species^{12,13} or the 2% rule in birds³⁵. The very large datasets generated using next-generation sequencing have considerable potential to facilitate more accurate dating by identifying nuclear markers that conform to the assumptions of the molecular clock but, curiously, fixed rates based on mtDNA data are still being used to calibrate such datasets when no suitable fossil calibration points exist³⁶. A possible solution may involve the identification of a suite of neutral markers that can be used to assess divergence between the species used in the original molecular dating studies, and 28S rRNA may be a suitable candidate.

Data availability

DNA sequences generated in this study were submitted to GenBank (COI accession numbers: MT331645–MT331814; 28S rRNA accession numbers: MT329760–MT330099).

Acknowledgements

A previous version of this article is available on bioRxiv: <https://doi.org/10.1101/2020.03.31.017764>.

References

- Kimura M: **The neutral theory of molecular evolution**. Cambridge: Cambridge University Press; 1983.
[Reference Source](#)
- Ballard JW, Kreitman M: **Is mitochondrial DNA a strictly neutral marker?** *Trends Ecol Evol*. 1995; **10**(12): 485–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Blier PU, Dufresne F, Burton RS: **Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation**. *Trends Genet*. 2001; **17**(7): 400–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ballard JW, Rand DM: **The population biology of mitochondrial DNA and its phylogenetic implications**. *Annu Rev Ecol Evol Syst*. 2005; **36**(1): 621–42.
[Publisher Full Text](#)
- Meiklejohn CD, Montooth KL, Rand DM: **Positive and negative selection on the mitochondrial genome**. *Trends Genet*. 2007; **23**(6): 259–63.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stewart JB, Freyer C, Elson JL, et al.: **Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease**. *Nat Rev Genet*. 2008; **9**(9): 657–62.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bazin E, Glémin S, Galtier N: **Population size does not influence mitochondrial genetic diversity in animals**. *Science*. 2006; **312**(5773): 570–2.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Teske PR, Golla TR, Sandoval-Castillo J, et al.: **Mitochondrial DNA is unsuitable to test for isolation by distance**. *Sci Rep*. 2018; **8**(1): 8448.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Soares P, Ermini L, Thomson N, et al.: **Correcting for purifying selection: an improved human mitochondrial molecular clock**. *Am J Hum Genet*. 2009; **84**(6): 740–59.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ho SY, Lanfear R, Bromham L, et al.: **Time-dependent rates of molecular evolution**. *Mol Ecol*. 2011; **20**(15): 3087–101.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Lawrie DS, Messer PW, Hershberg R, et al.: **Strong purifying selection at synonymous sites in *D. melanogaster***. *PLoS Genet*. 2013; **9**(5): e1003527.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Schubart CD, Diesel R, Hedges B: **Rapid evolution to terrestrial life in Jamaican crabs**. *Nature*. 1998; **393**: 363–5.
[Publisher Full Text](#)
- Knowlton N, Weigt LA: **New dates and new rates for divergence across the Isthmus of Panama**. *Proc Biol Sci*. 1998; **265**(1412): 2257.
[Publisher Full Text](#) | [Free Full Text](#)
- Lamb AM, Gan HM, Greening C, et al.: **Climate-driven mitochondrial selection: a test in Australian songbirds**. *Mol Ecol*. 2018; **27**(4): 898–918.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sun JT, Jin PY, Hoffmann AA, et al.: **Evolutionary divergence of mitochondrial genomes in two *Tetranychus* species distributed across different climates**. *Insect Mol Biol*. 2018; **27**(6): 698–709.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Matumba TG: **Genetics and thermal biology of littorinid snails of the genera *Afrolittorina*, *Echinolittorina* and *Littoraria* (Gastropoda: Littorinidae) from temperate, subtropical and tropical regions**. PhD thesis, Rhodes University, South Africa.
[Reference Source](#)
- Williams ST, Reid DG: **Speciation and diversity on tropical rocky shores: a global phylogeny of snails of the genus *Echinolittorina***. *Evolution*. 2004; **58**(10): 2227–2251.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Malaquias MAE, Reid DG: **Tethyan vicariance, relictualism and speciation: evidence from a global molecular phylogeny of the opisthobranch genus *Bulla***. *J Biogeogr*. 2009; **36**(9): 1760–1777.
[Publisher Full Text](#)
- Doyle J: **DNA protocols for plants**. In: *Molecular techniques in taxonomy*. 1991; 283–93.
[Publisher Full Text](#)
- Folmer O, Black M, Hoeh W, et al.: **DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates**. *Mol Mar Biol Biotechnol*. 1994; **3**(5): 294–9.
[PubMed Abstract](#)
- Littlewood DT, Curini-Galletti M, Herniou EA: **The interrelationships of Proseriata (Platyhelminthes: Seriata) tested with molecules and morphology**. *Mol Phylogenet Evol*. 2000; **16**(3): 449–66.
[PubMed Abstract](#) | [Publisher Full Text](#)

22. Williams ST, Reid DG, Littlewood DT: **A molecular phylogeny of the Littorininae (Gastropoda: Littorinidae): unequal evolutionary rates, morphological parallelism, and biogeography of the Southern Ocean.** *Mol Phylogenet Evol.* 2003; **28**(1): 60–86.
[PubMed Abstract](#) | [Publisher Full Text](#)
23. Kumar S, Stecher G, Tamura K: **MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.** *Mol Biol Evol.* 2016; **33**(7): 1870–4.
[PubMed Abstract](#) | [Publisher Full Text](#)
24. Stephens M, Smith NJ, Donnelly P: **A new statistical method for haplotype reconstruction from population data.** *Am J Hum Genet.* 2001; **68**(4): 978–89.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. Bandelt HJ, Forster P, Röhl A: **Median-joining networks for inferring intraspecific phylogenies.** *Mol Biol Evol.* 1999; **16**(1): 37–48.
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Leigh JW, Bryant D: **popart: full-feature software for haplotype network construction.** *Methods Ecol Evol.* 2015; **6**(9): 1110–6.
[Publisher Full Text](#)
27. Rogers AR, Harpending H: **Population growth makes waves in the distribution of pairwise genetic differences.** *Mol Biol Evol.* 1992; **9**(3): 552–69.
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Excoffier L, Lischer HEL: **Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows.** *Mol Ecol Resour.* 2010; **10**(3): 564–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
29. Stoeckle MY, Thaler DS: **DNA barcoding works in practice but not in (neutral) theory.** *PLoS One.* 2014; **9**(7): e100755.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Drummond AJ, Rambaut A, Shapiro B, *et al.*: **Bayesian coalescent inference of past population dynamics from molecular sequences.** *Mol Biol Evol.* 2005; **22**(5): 1185–92.
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Low VL, Norma-Rashid Y, Yusoff A, *et al.*: **Pleistocene demographic expansion and high gene flow in the globe skimmer dragonfly *Pantala flavescens* Fabricius (Odonata: Libellulidae) in Peninsular Malaysia.** *Zool Anz.* 2017; **266**: 23–7.
[Publisher Full Text](#)
32. Iván Pérez-Quiñonez C, Quiñonez-Velázquez C, García-Rodríguez FJ: **Genetic homogeneity of the Pacific thread herring (*Opisthonema libertate*) (Günther, 1867) in the Eastern Pacific, inferred from mtDNA sequences.** *Mitochondrial DNA A DNA Mapp Seq Anal.* 2019; **30**(3): 517–24.
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Diringer B, Pretell K, Avellan R, *et al.*: **Genetic structure, phylogeography, and demography of *Anadara tuberculosa* (Bivalvia) from East Pacific as revealed by mtDNA: implications to conservation.** *Ecol Evol.* 2019; **9**(8): 4392–402.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
34. Grant WS: **Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography.** *J Hered.* 2015; **106**(4): 333–46.
[PubMed Abstract](#) | [Publisher Full Text](#)
35. Shields GF, Wilson AC: **Calibration of mitochondrial DNA evolution in geese.** *J Mol Evol.* 1987; **24**(3): 212–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
36. Trucchi E, Gratton P, Whittington JD, *et al.*: **King penguin demography since the last glaciation inferred from genome-wide data.** *Proc Biol Sci.* 2014; **281**(1787): 20140528.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 2

Reviewer Report 04 September 2020

<https://doi.org/10.5256/f1000research.28975.r70500>

© 2020 Thaenkham U et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Abigail Hui En Chan

Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

Urusa Thaenkham

Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

We understand what the authors explained in the response to each comment. We know many limitations remarked in their responses. However, the readers who read this report in the future should realize that it was prepared and published under limitation conditions.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular systematics.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 12 August 2020

<https://doi.org/10.5256/f1000research.26080.r68386>

© 2020 Thaenkham U et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Abigail Hui En Chan

Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

Urusa Thaenkham

Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

This study aimed to explore the genetic variation of two common southern African snails of the genus *Afrolittorina*. The main finding of this study is that intraspecific variations of nuclear 28S rDNA sequences are higher than such value from the mitochondrial gene in the snail populations. However, many points need to concern before acceptance. I listed the major points for revising the manuscript.

1. Introduction part, the introduction part seems to focus on the selection process in the population to increase/decrease the genetic variation accumulation in the mt genome. The authors did not emphasize on the research questions; for example, Why did you want to study in two snails of the genus *Afrolittorina*? Why did you use the 28S rDNA as another genetic marker? What is the research question or hypothesis of this study? What was the reason behind deciding to study on *A. aficana* and *A. knysnaensis*? Finally, what is the main objective of this work? Whether you aimed to study the genetic differentiation between species of the snail based on different kinds of markers or study on population genetic, the content of this study could not support both objectives. For example, if the authors want to present on comparing the genetic variations that came out from the COI gene and 28S rDNA gene, the analysis used in this study was not suitable, the haplotype network analysis may not fit enough for the genetic variations you want to get. If the author wants to analyze the population genetically, the weak point is the numbers of the snails collected? And why the authors decided to collect from the various sites? What is the hypothesis behind your sampling?
2. I don't understand why the authors have to analyze the molecular clock by comparing different originated markers like COI from one of the protein-coding genes in the mitochondrial genome and LSU (RNA-specifying gene) from the nuclear genome. Of course, the rate of evolution acting on these two genomes is different.
3. If the authors considered in Figure 1, you would realize that 28S rDNA is not good enough for species discrimination between *A. aficana* and *A. knysnaensis* comparing with the COI gene. Only 2 nucleotide differences between those two snail species, while the nucleotide variation in the intra-specific level of *A. knysnaensis* is higher than that.
4. The authors didn't discuss the evidence in snails with regards to variation, which is the point that showed in the result.
5. There is no analysis to estimate either positive or negative selection, but the authors discussed it as the condition forced on the snail populations. It becomes over speculation.
6. The title of this study should mention snail species because the genetic variation patterns are various depending on the groups of organisms.

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

No

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

No

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

No

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

Not applicable

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

Yes

Are the conclusions drawn adequately supported by the results?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular systematics.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 19 Aug 2020

Peter Teske, University of Johannesburg, Auckland Park, South Africa

Reviewer Report: Abigail Hui En Chan and Urusa Thaenkham, Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand
APPROVED WITH RESERVATIONS

COMMENT: This study aimed to explore the genetic variation of two common southern African snails of the genus *Afrolittorina*. The main finding of this study is that intraspecific variations of nuclear 28S rDNA sequences are higher than such value from the mitochondrial gene in the snail populations.

RESPONSE: The statement that "intraspecific variations of nuclear 28S rDNA sequences are higher than such value from the mitochondrial gene" is not entirely correct, as in *A. africana*, COI has more haplotypes, whereas in *A. knysnaensis*, 28S has more haplotypes. We specifically state that "intra-specific genetic differentiation was comparatively low for this marker [i.e., COI], and similar to that of 28S."

COMMENT: However, many points need to concern before acceptance. I listed the major points for revising the manuscript. Introduction part, the introduction part seems to focus on the selection process in the population to increase/decrease the genetic variation accumulation in the mt genome. The authors did not emphasize on the research questions; for example, Why did you want to study in two snails of the genus *Afrolittorina*? Why did you use the 28S rDNA as another genetic marker? What is the research question or hypothesis of this study? What was the reason behind deciding to study on *A. africana* and *A. knysnaensis*? Finally, what is the main objective of this work?

RESPONSE: This article is a Brief Report, which has very strict word limits. For that reason, it is not possible to provide lengthy explanations for the issues raised. Briefly, the original purpose of the research was to determine whether each species is divided into regional populations (which was not found). The purpose was not to compare intraspecific variation between two markers with very different evolutionary rates, as that is an unexpected finding, and essentially a novel discovery that is reported in this paper. To answer the additional questions, we would like to refer the reviewers to the original PhD thesis on which this paper is based, which is now included among the references, and which is accessible via the following link:

http://vital.seals.ac.za:8080/vital/access/manager/Repository/vital:5588?site_name=GlobalView

COMMENT: Whether you aimed to study the genetic differentiation between species of the snail based on different kinds of markers or study on population genetic, the content of this study could not support both objectives. For example, if the authors want to present on comparing the genetic variations that came out from the COI gene and 28S rDNA gene, the analysis used in this study was not suitable, the haplotype network analysis may not fit enough for the genetic variations you want to get. If the author wants to analyze the population genetically, the weak point is the numbers of the snails collected?

RESPONSE: The aim of the study was not to analyse each population genetically. We assume here that the reviewers equate 'populations' with sampling sites, please see the following paper for an explanation why this is problematic: Waples & Gaggiotti (2006) *Molecular Ecology* 15:1419-1439. As neither species shows regional sub-structure (see also the PhD thesis mentioned earlier) and each species thus comprises a single 'population', it is possible to pool data from all sites for intraspecific genetic analysis. The number of samples collected (93 and 82) is actually large compared to other papers of this nature. For example, in their highly cited paper on the barcoding gap (whose detection depends on both intra- and inter-specific variation), Paulay & Meyer (2005) used an average number of 7.7 samples per cowrie species.

The fact that neither species of *Afrolittorina* is genetically subdivided throughout its range is clearly important, and was removed from an earlier version of this article so as not to exceed the word count for Brief Reports. We have decided to briefly mention it in the last paragraph of the Introduction: "Here, we explore this issue using mitochondrial and nuclear DNA sequence data from two common southern African snails of the genus *Afrolittorina* that show no spatial genetic structure throughout their ranges." We now also cite the PhD thesis for anyone interested in the biology of these snails.

COMMENT: And why the authors decided to collect from the various sites? What is the hypothesis behind your sampling?

RESPONSE: The original purpose of collecting samples from multiple sites that span much of each species' range was to determine whether or not these species are genetically subdivided into regional populations, which was not found. However, this is clearly not relevant to the present study, which focuses on intraspecific variation, not genetic structure. Obtaining large numbers of samples from a few sites would be problematic as one cannot rule out that additional, genetically distinct 'populations' exist that were not sampled. Please note that Table 1 is not particularly important to understand this paper, and for that reason should not be the focus of a lengthy discussion related to sampling. We originally had it in an Extended Data section, but the editor requested us to move it to the main text.

COMMENT: I don't understand why the authors have to analyze the molecular clock by comparing different originated markers like COI from one of the protein-coding genes in the mitochondrial genome and LSU (RNA-specifying gene) from the nuclear genome. Of course, the rate of evolution acting on these two genomes is different.

RESPONSE: The difference in evolutionary rates between these two markers is crucial to highlight the potential role of variation-reducing selection in COI. Without understanding this, it is not possible to understand the take-home message of this paper. We essentially point out that the known rate differences between these two markers manifest at the inter-specific level. The novel finding of this study is that intra-specific levels of variation are surprisingly similar. To clarify, the statement "Of course, the rate of evolution acting on these two genomes is different" is only true between species, not within species.

COMMENT: If the authors considered in Figure 1, you would realize that 28S rDNA is not good enough for species discrimination between *A. aficana* and *A. knysnaensis* comparing with the COI gene. Only 2 nucleotide differences between those two snail species, while the nucleotide variation in the intra-specific level of *A. knysnaensis* is higher than that.

RESPONSE: The purpose of this paper is not to suggest that 28S is a marker that is suitable for species discrimination, particularly when compared to COI. As we state in the first sentence of the Discussion, "The usefulness of the mtDNA COI gene to uncover overlooked biodiversity is undisputed because of the marker's tendency to have a well-defined barcoding gap, as was also found here." It is not clear what the reviewers mean by "not good enough", the two genetic clusters are clearly distinct, although inter-specific divergence is clearly much lower. The likely reason for this is that 28S evolves in a more clock-like fashion while COI is strongly affected by selective forces, and this is another important point to understand, as a marker that does not evolve in a clock-like fashion should not be used for intraspecific molecular dating. Please see the second paragraph in the Discussion: "In the present study, we have highlighted a largely unexplored problem that likely arises from selection effects in mtDNA data: the fact that demographic events using gene regions under purifying selection are dated using molecular clock calibrations affected by variation-increasing inter-specific divergence." Given the slow evolutionary rate of 28S, minimal differentiation between species is expected. What is unexpected are the similar levels of intraspecific variation (see title).

COMMENT: The authors didn't discuss the evidence in snails with regards to variation, which is the point that showed in the result.

RESPONSE: The results report the simplest and most intuitive means of describing intraspecific variation: the number of haplotypes, and the maximum number of mutations between them. Several sentences in the Discussion deal with this, e.g. first paragraph: "In contrast, there was comparatively little genetic variation at the intraspecific level for both markers, which is likely due to the commonly reported strong purifying selection acting upon the COI gene."

COMMENT: There is no analysis to estimate either positive or negative selection, but the authors discussed it as the condition forced on the snail populations. It becomes over speculation.

RESPONSE: The paper originally included a test for selection, but as it was not informative

and exceeded the word limit for Brief Reports, it was removed. We used a McDonald-Kreitman test (McDonald & Kreitman 1991) in MKT (Egea et al. 2008) to test for selection in COI. Out of 58 mutations for the whole dataset, only one was non-synonymous. Although the test had the highest possible proportion of adaptive substitutions (alpha) of 1.0, it was non-significant ($P = 0.57$), supposedly because of the lack of non-synonymous polymorphism.

Further, recent genomic evidence indicates that tests comparing synonymous and non-synonymous mutations are not actually conclusive about whether or not selection has taken place, because synonymous sites may be affected by strong selection even though no amino acid changes have taken place (Lawrie DS et al. (2013) Strong Purifying Selection at Synonymous Sites in *D. melanogaster*. PLoS Genet 9(5): e1003527). Again, a detailed discussion of these issues would far exceed the word limit, and would not add much to the study. The simple dataset used here is clearly not sufficiently informative to conclude that purifying selection is present, and for that reason we have changed the earlier title "Purifying selection can reduce intraspecific mitochondrial gene variation to that of nuclear rRNA" (<https://www.biorxiv.org/content/10.1101/2020.03.31.017764v1>) to the present one, which merely focuses on intraspecific variation rather than selection.

We nonetheless do not think that mentioning selection in this context is over-speculative. There is strong evidence that a finding of barcoding gaps in hundreds of studies is the result of selective forces, and we cite several important papers in this context, including two at the end of the following sentence in the Discussion: "In contrast, there was comparatively little genetic variation at the intraspecific level for both markers, which is likely due to the commonly reported strong purifying selection acting upon the COI gene." In terms of finding a COI barcoding gap, our study is merely an additional example, but the finding that intraspecific variation-reducing selection can be so significant that COI shows levels of variation similar to a marker that evolves much more slowly is novel.

COMMENT: The title of this study should mention snail species because the genetic variation patterns are various depending on the groups of organisms.

RESPONSE: We respectfully disagree – the practice of including taxonomic information in the title is standard procedure in taxonomy journals, but it is undesirable in a multidisciplinary journal such as F1000Research. In general, it is well known that shorter titles that report results are more likely to get accessed (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351256/>). In the present case, the snail species, their geographic structure, their biology etc. are far less important than are the two genetic markers, and anyone who does not work on littorinid snails may consider the paper irrelevant to their own work if species names are included in the title. However, the finding of this study is likely applicable to a wide range of organisms given the mounting evidence for variation-reducing selection in mtDNA, and giving the impression that it is unique to two rocky shore snails would be an obstacle to further investigation into this issue. To provide an example, had Stoeckle & Thaler 2014 mentioned in the title of their article "DNA barcoding works in practice but not in (neutral) theory" that these findings are based entirely on birds, we would probably not have cited their paper, even though it is highly relevant. For that reason, we believe that it should be sufficient to mention "rocky shore snails" in the Abstract.

Competing Interests: None.

Reviewer Report 20 July 2020

<https://doi.org/10.5256/f1000research.26080.r66088>

© 2020 Briscoe A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Andrew G. Briscoe

Natural History Museum, London, UK

This is an interesting manuscript addressing the methodological issues with the continued trend of using mitochondrial DNA sequence data to infer genetic variation and population structure. The results are based on a very limited data set, however, they are only intended to further illustrate the issues highlighted by the authors. Apart from that, the study does not require much improvement. It is well written and structured and the results support the conclusions and therefore suggest indexing the article once the following issues have been considered and adequately resolved;

The authors claim that this is the first study to use 28S ribosomal RNA at this taxonomic level. Do they mean in snails, as this gene has been used extensively for phylogenetics and some population level analysis? More info is needed.

Could the authors include a reference when stating that there is an order of magnitude of difference between the mutation rates of the two markers used in the study?

Reference 29 is cited as an example of the continued use of mtDNA for exploring demography, however, that study uses the mitochondrial control region rather than sequence data Protein Coding Genes (PCGs). Is there evidence that the mutation rates of this mitochondrial locus is under the same kind of diversifying selection as PCGs? If not, then the authors should consider changing or removing this reference.

In the discussion the authors refer to the different thermal environments of the two snail species used in the study. It would be useful if the authors could elaborate on this as very little background about the species being studied is provided.

The authors state that they are assuming 28S evolves in a clock-like manner. Can they provide a references/evidence for this as it is quite a bold statement, which has implications on some of the conclusions being drawn.

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

Yes

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?

Yes

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

Yes

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

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular biology, mitogenomics, phylogenetics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 19 Aug 2020

Peter Teske, University of Johannesburg, Auckland Park, South Africa

This is an interesting manuscript addressing the methodological issues with the continued trend of using mitochondrial DNA sequence data to infer genetic variation and population structure. The results are based on a very limited data set, however, they are only intended to further illustrate the issues highlighted by the authors. Apart from that, the study does not require much improvement. It is well written and structured and the results support the conclusions and therefore suggest indexing the article once the following issues have been considered and adequately resolved;

COMMENT: The authors claim that this is the first study to use 28S ribosomal RNA at this taxonomic level. Do they mean in snails, as this gene has been used extensively for phylogenetics and some population level analysis? More info is needed.

RESPONSE: We have conducted an extensive literature review on this issue, and have not found any comparable example. Even though intraspecific variation in 18S or 28S rRNA has been reported in various invertebrate taxa, upon closer inspection it becomes evident that this variation is either located in the more rapidly evolving ITS regions rather than in the rRNA, or the species studied have geographically distinct populations that may constitute cryptic species. However, as we cannot rule out the possibility that comparable examples exist, we have changed the sentence to: "Nuclear rRNA is not usually used to study intraspecific variation in species that are not spatially structured, presumably because this marker is assumed to evolve so slowly that it is more suitable for phylogenetics."

COMMENT: Could the authors include a reference when stating that there is an order of magnitude of difference between the mutation rates of the two markers used in the study?

RESPONSE: The paper already includes several references that have estimated mutation rates specifically for marine snails. Please see Table 2, where the 28S rates range from 0.01-0.05 %/myr, whereas the COI rates range from 0.26-0.50 %/myr. We now cite both references in the last paragraph of the Introduction.

COMMENT: Reference 29 is cited as an example of the continued use of mtDNA for exploring demography, however, that study uses the mitochondrial control region rather than sequence data Protein Coding Genes (PCGs). Is there evidence that the mutation rates of this mitochondrial locus is under the same kind of diversifying selection as PCGs? If not, then the authors should consider changing or removing this reference.

RESPONSE: This is a valid point, and while a signature of diversifying selection in the control region is likely because this marker is linked to the PCGs by virtue of being part of the same genome, it is clearly not under selection itself. For that reason, we have decided to remove this reference.

COMMENT: In the discussion the authors refer to the different thermal environments of the two snail species used in the study. It would be useful if the authors could elaborate on this as very little background about the species being studied is provided.

RESPONSE: To comply with word count requirements for Brief Reports, we have removed more detailed descriptions of the species' distribution ranges, morphology and physiology. In response to similar queries from the other reviewers, we now cite the PhD thesis on which the present study is based. This thesis not only provides information on the temperature-defined marine biogeographical provinces in which these species occur, but it also shows evidence for adaptive differences assessed using physiological experiments.

COMMENT: The authors state that they are assuming 28S evolves in a clock-like manner. Can they provide a references/evidence for this as it is quite a bold statement, which has implications on some of the conclusions being drawn.

RESPONSE: We have found no specific information on this, as the issue of selection on the molecular clock has primarily received attention in the context of mtDNA. However, we believe that the considerable difference in terms of genetic differentiation between species justifies such as statement. There is now strong evidence that the barcoding gap in COI is a result of selective forces, and the comparatively low level of genetic divergence based on 28S implies that similar selective forces are not acting on this marker. We have, however, made the statement in the Abstract less bold and have changed it to: "Assuming that 28S evolution is more clock-like...".

Competing Interests: None

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

F1000Research