GENOME NOTE

The first genome sequence of *Anopheles squamous* from Macha, Zambia [version 1; peer review: 2 approved]

Valerie T. Nguyen[1], Travis C. Collier[2], Sangwoo Seok[1], Xiaodi Wang[1], Monicah M. Mburu[3], Limonty Simubali[3], Mary E. Gebhardt[4], Douglas E. Norris[4], Yoosook Lee[1]

[1]Florida Medical Entomology Laboratory, Department of Entomology and Nematology, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida, 32962, USA
[2]Independent researcher, Vero Beach, FL, 32962, USA
[3]Macha Research Trust, Macha, Southern Province, Zambia
[4]The W. Harry Feinstone Department of Molecular Microbiology and Immunology, The Johns Hopkins Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, 21205, USA

v1 **First published:** 24 Mar 2023, **12**:330
https://doi.org/10.12688/f1000research.130734.1
Latest published: 24 Mar 2023, **12**:330
https://doi.org/10.12688/f1000research.130734.1

Abstract

Despite efforts to minimize the impacts of malaria and reduce the number of primary vectors, malaria has yet to be eliminated in Zambia. Understudied vector species may perpetuate malaria transmission in pre-elimination settings. *Anopheles squamosus* is one of the most abundantly caught mosquito species in southern Zambia and has previously been found with *Plasmodium falciparum* sporozoites, a causal agent of human malaria. This species may be a critical vector of malaria transmission, however, there is a lack of genetic information available for *An. squamosus*. We report the first genome data and the first complete mitogenome (Mt) sequence of *An. squamosus*. The sequence was extracted from one individual mosquito from the Chidakwa area in Macha, Zambia. The raw reads were obtained using Illumina Novaseq 6000 and assembled through NOVOplasty alignment with related species. The length of the *An. squamosus* Mt was 15,351 bp, with 77.9 % AT content. The closest match to the whole mitochondrial genome in the phylogenetic tree is the African malaria mosquito, *Anopheles gambiae*. Its genome data is available through National Center for Biotechnology Information (NCBI) Sequencing Reads Archive (SRA) with accession number SRR22114392. The mitochondrial genome was deposited in NCBI GenBank with the accession number OP776919. The ITS2 containing contig sequence was deposited in GenBank with the accession number OQ241725. Mitogenome annotation and a phylogenetic tree with related *Anopheles* mosquito species are provided.

Keywords

Anopheles squamosus, understudied malaria vector, Africa, Zambia

Open Peer Review

Approval Status ✓ ✓

	1	2
version 1 24 Mar 2023	✓ view	✓ view

1. **Megan A. Riddin**, University of Pretoria, Pretoria, South Africa

2. **Tara Roth**, San Mateo County Mosquito and Vector Control District, Burlingame, USA

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

This article is included in the Genomics and Genetics gateway.

Corresponding author: Yoosook Lee (yoosook.lee@ufl.edu)

Author roles: Nguyen VT: Conceptualization, Data Curation, Formal Analysis, Investigation, Writing – Original Draft Preparation; **Collier TC**: Formal Analysis, Writing – Review & Editing; **Seok S**: Formal Analysis, Investigation, Writing – Review & Editing; **Wang X**: Formal Analysis, Investigation, Writing – Review & Editing; **Mburu MM**: Data Curation, Investigation, Resources, Writing – Review & Editing; **Simubali L**: Data Curation, Investigation, Writing – Review & Editing; **Gebhardt ME**: Data Curation, Investigation, Resources, Writing – Review & Editing; **Norris DE**: Conceptualization, Resources, Writing – Review & Editing; **Lee Y**: Conceptualization, Resources, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This study was supported by University of Florida College of Agricultural and Life Sciences Dean's Award to VTN, University of Florida Department of Entomology and Nematology Matching Assistantship to SS, University of Florida International Center Global Fellows Program Award to YL, NIH T32 Training Grant (T32 AI-138953) support to MEG, NIH support through the International Centers of Excellence for Malaria Research (2U19AI089680) for DEN, MEG, MMM and LS, and support to DEN and MEG from the Bloomberg Philanthropies and the Johns Hopkins Malaria Research Institute.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Nguyen VT, Collier TC, Seok S *et al.* **The first genome sequence of *Anopheles squamous* from Macha, Zambia [version 1; peer review: 2 approved]** F1000Research 2023, **12**:330 https://doi.org/10.12688/f1000research.130734.1

First published: 24 Mar 2023, **12**:330 https://doi.org/10.12688/f1000research.130734.1

Introduction

Anopheles squamosus (Theobald, 1901; Figure 1) can be found across Africa[1] and is of particular relevance to public health due to its implication in the spread of residual malaria cases. *Anopheles squamosus* is one of the most abundantly caught anopheline species in malaria vector surveillance studies in southern Zambia. However, it is understudied species because of its exophilic and zoophilic behaviours.[2,3] Though they are predominantly associated as a zoophilic species, they have been discovered to have high anthropophily in southern Zambia.[4] Additionally, there has been the detection of *Plasmodium falciparum* sporozoite and DNA, a causal agent of human malaria, in *An. squamosus*.[2,5]

Unfortunately, there are two key barriers to pursuing the rigorous investigation of the epidemiologically important traits of this vector, such as host choice, biting behaviours, and dispersal capacity. First, *An. squamosus* is morphologically indistinguishable from *An. cydippis* at the adult stage. Although they are morphologically distinct as larvae, larvae are often difficult to locate in abundance. There are numerous examples of sympatric *Anopheles* sibling species expressing drastic differences in insecticide resistance[6] or host choice.[7,8] These differences make species confirmation critical to assessing and mitigating malaria transmission risk. Second, there is limited genetic information (173 sequences total in GenBank as of August 2022) for *An. squamosus*, most (N=166; 96%), are partial sequences of the mitochondrial cytochrome c oxidase subunit I (COI) gene. ITS2 sequences are better at differentiating species within a complex than COI sequences[9] but existing ITS2 primers do not typically work on *An. squamosus* and the absence of sequence data for this region prohibit the design of functional diagnostic PCR primers.

To overcome these two barriers and advance investigative efforts aimed at this widespread, yet neglected malaria vector, we carried out the first Illumina high-throughput sequencing of this species.

Methods

Data collection

The *An. squamosus* sample used for the genome sequencing was collected in Chidakwa near Macha, Zambia (utm-x: 0478202, utm-y: 8184394) using a CDC light trap placed outdoors near a goat pen. Samples were frozen after collection at –20°C until DNA extraction. DNA was extracted using a magnetic bead-based protocol optimized for mosquito DNA for Next-generation sequencing.[10] The head and thorax were dissected from the sample and hydrated in nuclease-free water for 1 hour at 4°C. Tissues were then removed from the water and homogenised in a mixture of 2 µL Proteinase K (100 mg/mL) and 98 µL PK Buffer in a 1.5 mL Eppendorf microcentrifuge tube (add company name), followed by incubation at 56°C for 2 hours. The lysate was transferred to a new 1.5 mL microcentrifuge tube and mixed with a MagAttract Mix consisting of 100 µL isopropanol, 100 µL Buffer AL, and 15 µL MagAttract Suspension G (Qiagen, Hilden, Germany). The mixture was incubated at room temperature for 10 minutes and occasionally vortexed to ensure that the magnetic beads were evenly dispersed. The microcentrifuge tube containing the lysate was then moved to a magnetic bead separator until the liquid appeared clear. After a series of ethanol washes of magnetic beads, DNA was eluted from the beads with 100 µL AE Buffer and stored at −20°C until library preparation. The library preparation was

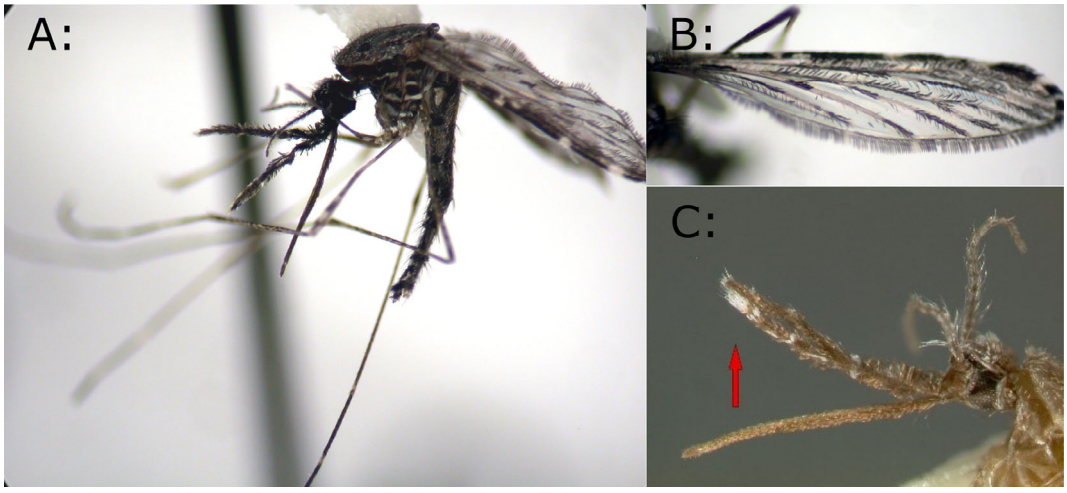

Figure 1. A: *Anopheles squamosus* image. B: *An. squamosus* wing. C: Head image of *An. squamosus*, one of the features used for species identification. A and B have been reproduced with permission from Dr. Rebekah Kading (Colorado State University).[21] C has been reproduced with permission from Walter Reed Biosystematics Unit (WRBU).[22]

completed using the QIAseq FX UDI kit (Qiagen, Hilden, Germany) using 20 ng genomic DNA as input for the protocol as previously described.[11]

Data analysis

Raw sequencing reads were trimmed using fastp (RRID:SCR_016962) version 0.20.1.[12] Mitogenome (Mt) contig was assembled using NOVOPlasty (RRID:SCR_017335) version 4.3.1.[13] Automatic annotation of mitogenome was conducted with the MITOS website[14] using the invertebrate genetic code for mitochondria under default settings. Some automatic annotations were not consistent with typical *Anopheles* mitochondrial gene start and/or end positions. Manual adjustments were made to inconsistent automatic annotations by shifting the start and end positions to match existing *Anopheles* mitochondrial gene annotations found in GenBank. Annotation information was also deposited to the GenBank with the genome sequence. The full genomic map is provided in Figure 2.

Phylogenetic analysis was conducted using the mitogenome sequences of seven *Anopheles* species and one *Aedes* species as an outgroup in**.** The Jukes-Cantor model was used to calculate the pairwise genetic distances and the neighbour-joining method was used to build the phylogenetic tree in Geneious Prime (RRID:SCR_010519) 2022.02 (Biomatters, Auckland, New Zealand)[15] (free alternative, AliView).

Draft genome assembly was conducted using MaSuRCA (RRID:SCR_010691) version 4.0.9[16] in order to find a contig containing Internal transcribed spacer 2 (ITS2) sequence. Basic local alignment search tool (BLAST) (RRID: SCR_004870) was used for the resulting contigs to locate contigs with highest similarity with only *An. squamosus* ITS2 sequence available on GenBank (accession number MK592071).

Results

We yielded 105 million reads from sequencing a single *An. squamosus* sample. Of these, 238,740 reads were used to assemble mitochondrial genome. Draft genome assembly using MaSuRCA produced 58,252 scaffolds with the total size of scaffolds of 350Mbp. N50 scaffold length was 21,439bp. Among these contigs, we identified one contig containing ITS2 region (GenBank accession number OQ241725), which was 1,223 bp long.

The length of the *An. squamosus* Mt (GenBank accession number OP776919[23]) was 15,351 bp and the percentage A+T was 77.9% (Figure 2). The average A+T percentage of eight other anopheline species was 77.7% (±0.61 SD). The length of this mitochondrial genome was a similar length to other anopheline species that have been deposited in GenBank, with the average of the eight species compared in this analysis being 15,363 bp. The content for this mitochondrial genome includes two ribosomal RNAs, 22 transfer RNAs, and 35 protein-coding genes. The cytochrome c oxidase I (COI) fragment spanning 1,462-2,132 bp of *An. squamosus* sequence had 97.7% (±4.27 SD, N=9) similarity to the COI sequence of *An. squamosus* deposited in GenBank.

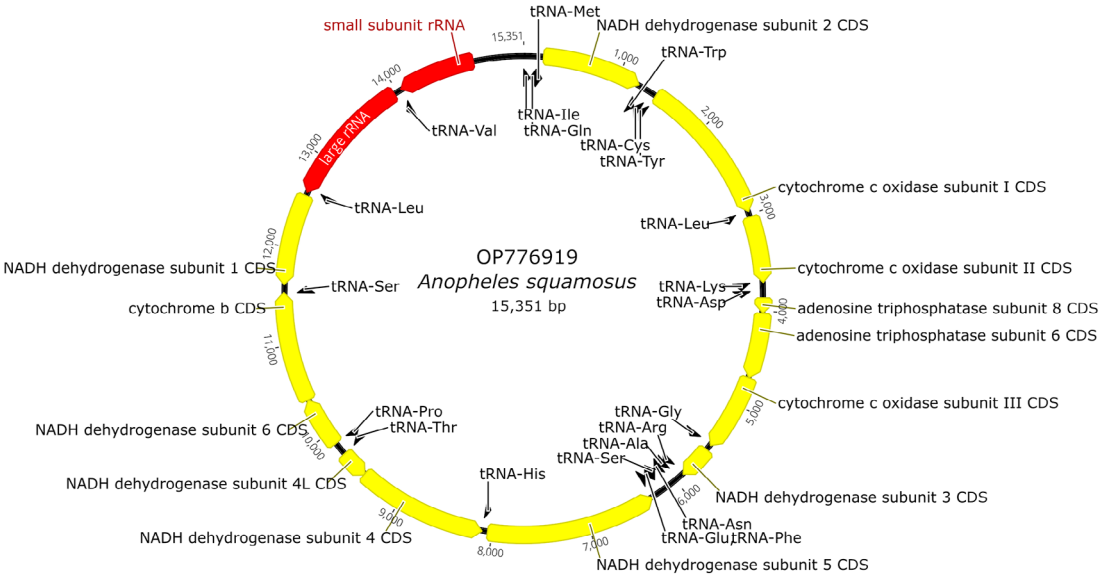

Figure 2. Mitogenome map of *An. squamosus* including annotated genes.

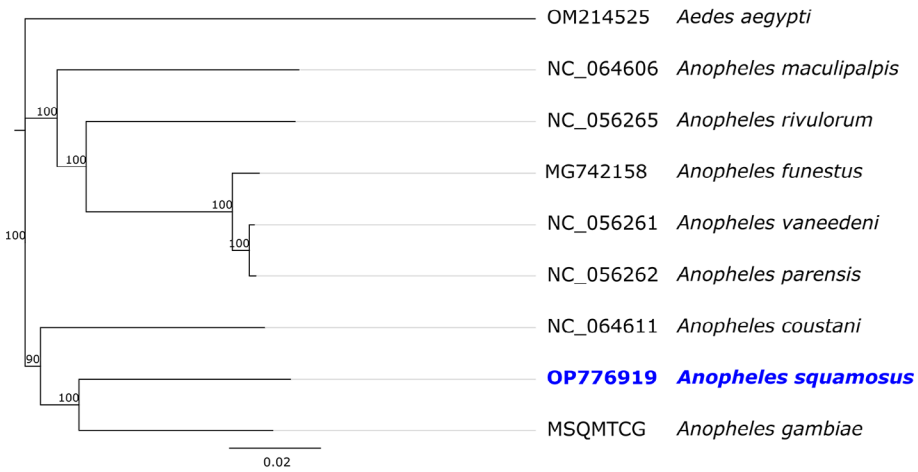

Figure 3. Phylogenetic tree based on mitogenome sequences of *An. squamosus* and its related mosquitoes. Species names are provided next to the GenBank accession numbers. Numbers at nodes indicate bootstrap values out of 100 replicates. *Aedes aegypti* was considered as an outgroup. The scale bar indicates relative nucleotide difference (0.02=2% nucleotide difference).

In the phylogenetic analysis (Figure 3), the closest match to the whole mitogenome sequence of *An. squamosus* was the African malaria mosquito *An. gambiae* (GenBank accession number L20934), with 91.5% sequence similarity. This comparable sample was identified as *An. gambiae* and published in 1993 before *An. gambiae* were separated into two species: *An. gambiae* and *An. coluzzii.*[17] Nevertheless, previous studies suggest that mitogenome sequence alone is not sufficient to distinguish *An. gambiae* s.s. from *An. coluzzii.*[18,19]

This study provides a critical genomic resource for research of this understudied malaria vector. Our short reads sequencing data was not sufficient to assemble high-quality reference genome and revealed the need for alternative long-read sequencing technology for a high-quality genome assembly. However, we provided a key ITS2 region data that researchers can develop a low-cost molecular diagnostic assay to identify species. Currently available ITS2 primers for anopheline species identification typically does not produce a PCR amplicon, which is one of the major roadblocks in carrying out surveillance and research of this species. We identified the ITS2-containing contig (GenBank Accession number OQ241725) that could be used for new primer design that would amply the ITS2 fragments more reliably for *An. squamosus.* Our genome sequence data could be used for further variant identification once high-quality reference genome become available for *An. squamosus.* The mitogenome sequence could also be used to identify phylogenetic relationship within and between related species and infer gene flow/dispersal.[9,20]

Ethical considerations

The study involves collection of mosquito specimen near goat pens within individual households in Chidakwa, Zambia as part of the project that had been approved by National Health Research Authority, Zambia: Approval No: NHRA00016/18/08/2021.

Data availability

Underlying data

GenBank: *Anopheles squamosus* mitochondrion, complete genome. Accession number OP776919; https://identifiers.org/ncbi/insdc:OP776919.[23]

BioProject: Complete mitogenome sequence of *Anopheles squamosus* from Macha, Zambia. Accession number PRJNA896235; https://identifiers.org/bioproject:PRJNA896235.[24]

SRA: DNA-Seq of mosquito *Anopheles squamosus.* Accession number SRR22114392; https://identifiers.org/insdc.sra:SRR22114392.[25]

BioSample: *Anopheles squamosus* isolate As22MACHA01. Accession number SAMN31538381; https://identifiers.org/biosample:SAMN31538381.[26]

Acknowledgements

We thank UF ICBR for providing sequencing services. We appreciate the support from Dr. Edgar Simulundu from Macha Research Trust toward our project.

References

1. Theobald FV: *A monograph of the Culicidae, or mosquitoes.* London: British museum (Nat. hist.). Department of Zoology; 1901; 256.
2. Stevenson JC, Simubali L, Mbambara S, *et al.*: **Detection of *Plasmodium falciparum* Infection in *Anopheles squamosus* (Diptera: Culicidae) in an Area Targeted for Malaria Elimination, Southern Zambia.** *J. Med. Entomol.* 2016 Nov; **53**(6): 1482–1487.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Hoffman JE, Ciubotariu SL, Mudenda T, *et al.*: **Phylogenetic Complexity of Morphologically Identified *Anopheles squamosus* in Southern Zambia.** *Insects.* 2021 Feb 8; **12**(2).
PubMed Abstract | Publisher Full Text | Free Full Text
4. Fornadel CM, Norris LC, Franco V, *et al.*: **Unexpected anthropophily in the potential secondary malaria vectors *Anopheles coustani* s.l. and *Anopheles squamosus* in Macha, Zambia.** *Vector Borne Zoonotic Dis.* 2011 Aug; **11**(8): 1173–1179.
PubMed Abstract | Publisher Full Text | Free Full Text
5. Stevenson JC, Norris DE: **Implicating cryptic and novel anophelines as malaria vectors in Africa.** *Insects.* 2017; **8**(1): 1–18.
6. Main BJ, Lee Y, Collier TC, *et al.*: **Complex genome evolution in *Anopheles coluzzii* associated with increased insecticide usage in Mali.** *Mol. Ecol.* 2015 Oct; **24**(20): 5145–5157.
PubMed Abstract | Publisher Full Text | Free Full Text
7. Kent RJ, Thuma PE, Mharakurwa S, *et al.*: **Seasonality, blood feeding behavior, and transmission of *Plasmodium falciparum* by *Anopheles arabiensis* after an extended drought in southern Zambia.** *Am. J. Trop. Med. Hyg.* 2007 Feb; **76**(2): 267–274.
PubMed Abstract | Publisher Full Text | Free Full Text
8. Main BJ, Lee Y, Ferguson HM, *et al.*: **The genetic basis of host preference and resting behavior in the major African malaria vector, *Anopheles arabiensis.*** *PLoS Genet.* 2016 Sep; **12**(9): e1006303.
PubMed Abstract | Publisher Full Text | Free Full Text
9. Jones CM, Ciubotariu II, Muleba M, *et al.*: **Multiple novel clades of anopheline mosquitoes caught outdoors in northern Zambia.** *Frontiers Trop. Dis.* 2021; **2**: 780664.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Chen T, Vorsino AE, Kosinski KJ, *et al.*: **A magnetic-bead-based mosquito DNA extraction protocol for Next-Generation sequencing.** *J. Vis. Exp.* 2021; **170**: e62354.
Publisher Full Text
11. Kelly ET, Mack LK, Campos M, *et al.*: **Evidence of local extinction and reintroduction of *Aedes aegypti* in Exeter, California.** *Front. Trop. Dis.* 2021 Jul; **2**(8).
Publisher Full Text
12. Chen S, Zhou Y, Chen Y, *et al.*: **fastp: an ultra-fast all-in-one FASTQ preprocessor.** *Bioinformatics.* 2018 Sep 1; **34**(17): i884–i890.
PubMed Abstract | Publisher Full Text | Free Full Text
13. Dierckxsens N, Mardulyn P, Smits G: **NOVOPlasty: de novo assembly of organelle genomes from whole genome data.** *Nucleic Acids Res.* 2017 Feb 28; **45**(4): e18.
PubMed Abstract | Publisher Full Text
14. Bernt M, Donath A, Juhling F, *et al.*: **MITOS: improved de novo metazoan mitochondrial genome annotation.** *Mol. Phylogenet. Evol.* 2013 Nov; **69**(2): 313–319.
PubMed Abstract | Publisher Full Text
15. Kearse M, Moir R, Wilson A, *et al.*: **Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.** *Bioinformatics.* 2012 Jun 15; **28**(12): 1647–1649.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Zimin AV, Marçais G, Puiu D, *et al.*: **The MaSuRCA genome assembler.** *Bioinformatics.* 2013 Nov 1; **29**(21): 2669–2677.
PubMed Abstract | Publisher Full Text | Free Full Text
17. Coetzee M, Hunt RH, Wilkerson RC, *et al.*: ***Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex.** *Zootaxa.* 2013; **3619**(2): 246–274.
Publisher Full Text
18. Hanemaaijer MJ, Houston PD, Collier TC, *et al.*: **Mitochondrial genomes of *Anopheles arabiensis*, *An. gambiae* and *An. coluzzii* show no clear species division.** *F1000Res.* 2018; **7**: 347.
Publisher Full Text
19. Fontaine MC, Pease JB, Steele A, *et al.*: **Mosquito genomics. Extensive introgression in a malaria vector species complex revealed by phylogenomics.** *Science.* 2015 Jan 2; **347**(6217): 1258524.
Publisher Full Text
20. Campos M, Patel N, Marshall C, *et al.*: **Population genetics of *Anopheles pretoriensis* in Grande Comore Island.** *Insects.* 2023; **14**(1): 14.
PubMed Abstract | Publisher Full Text | Free Full Text
21. Kent RJ: **The Mosquitoes of Macha, Zambia.** *The third International Malaria Research Conference.* Baltimore, MD, USA: 2006 [cited 2022 Nov 8].
Reference Source
22. Walter Reed Biosystematics Unit (WRBU): **Anopheles squamosus Theobald.** 1901. [cited 2022 Nov 3].
Reference Source
23. Nguyen VT, Gebhardt ME, Mburu MM, *et al.*: *Anopheles squamosus* mitochondrion, complete genome. [Dataset]. *GenBank.* 2023.
Reference Source
24. University of Florida: Complete mitogenome sequence of *Anopheles squamosus* from Macha, Zambia. [Dataset]. *BioProject.* 2022.
Reference Source
25. University of Florida: DNA-Seq of mosquito *Anopheles squamosus.* [Dataset]. *SRA.* 2022.
Reference Source
26. University of Florida: *Anopheles squamosus* isolate As22MACHA01. [Dataset]. *BioSample.* 2022.
Reference Source

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 11 October 2023

https://doi.org/10.5256/f1000research.143510.r200578

✔ **Tara Roth**
San Mateo County Mosquito and Vector Control District, Burlingame, California, USA

This article details the production of the An. squamosus full genome which is intended to provide a useful reference for future researchers to investigate the role this species may play in the ecologic cycle of malaria in Zambia. The article stresses that mitogenome sequences (predominantly CO1) are not sufficient to differentiate closely related species which is a consistent issue across Culicidae. They also stress that primers for the alternative region, ITS2, do not amplify An. squamosus. The article is well written, however there are a few typos that should be addressed. Under Methods -> Data Collection, the line "Eppendorf microcentrifuge tube (add company name)" should be fixed. The final paragraph of the results section could be tightened up a bit - there are a few minor grammatical errors (eg: "The mitogenome sequence could also be used to identify phylogenetic relationship within...")

Are the rationale for sequencing the genome and the species significance clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of the sequencing and extraction, software used, and materials provided to allow replication by others?
Yes

Are the datasets clearly presented in a usable and accessible format, and the assembly and annotation available in an appropriate subject-specific repository?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a vector ecologist with a specialization in mosquitoes and ticks. I am

currently working on a project addressing the distribution of sister Anopheline species An. hermsi and An. freeborni in California and thus am familiar with the regions (CO1 and ITS2) being discussed in this paper

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.

Reviewer Report 05 April 2023

https://doi.org/10.5256/f1000research.143510.r167805

Megan A. Riddin
UP Institute for Sustainable Malaria Control (UP ISMC), Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

This research presented the first genome data and complete mitogenome sequence of *Anopheles squamosus*. *Anopheles squamosus* is commonly caught in abundance in Zambia and has been recently implicated in malaria transmission through the detection of P. falciparum circumsporozoites and through PCR. This research utilised a single *An. squamosus* specimen from Chidakwa, Zambia, which was collected through a CDC light trap placed outside near a goat pen. DNA was extracted from the specimen following a magnetic bead-based protocol specific for mosquito DNA for next generation sequencing. The library was then prepared using the eluted DNA and a standardised preparation kit (QIAseq FX UDI Kit). The sequencing reads were trimmed and mitogenome assembled using NOVOPlasty, and automatic annotation completed through the MITOS website. Phylogenetic analyses were performed with an *Aedes* species outgroup in Geneious Prime. Final draft genome assembly was conducted using MaSuRCA.
The methods resulted in 105 million reads and one contig was identified as containing an ITS2 region. The mitochondrial genome was assembled and resulted in 15351 bp with 77% A+T percentage. Phylogenetic analyses identified the closest mitogenome sequence match to *An. gambiae* which was sequenced prior to separation of the complex into sibling species.

The research enabled identification of key ITS2 region data for development of low-cost molecular diagnostic assays, despite the short reads not being sufficient for high-quality reference genome assembly. The outputs further included the identification of the phylogenetic relationships to related species through assembled mitogenome sequence.

This study presented vital data on an understudied species that has recently been implicated in malaria transmission in Zambia. Malaria has experienced recent upsurges in Zambia and although the primary vectors are thought to be *An. funestus*, *An. gambiae* and *An. arabiensis*, the detection of higher abundance of *An. squamosus* and *An. coustani*, as well as their unexpected anthopophillic behaviour has lent support to their role in transmission. The recent detection of both *P. falciparum circumsporozoite* in *An. squamosus* has highlighted the necessity to study this species. The use of

full genome sequencing is highly beneficial not only for phylogenetic analysis but further for development of rapid molecular identification techniques as was exhibited in this study. This study was able to provide data that will assist towards classification and identification of the species to better understand ecology, epidemiology and therefore role of this species in malaria transmission in Zambia.

The Introduction and rationale of the study is clearly explained and may benefit from some clarification on the chromosomal types or mention of possible clades that have been eluded to. However, as such work and suggestions are unpublished, it would be very interesting to see how the authors progress with this research towards understanding the phylogeny of the understudied species.

The Methods and Results are clear and concise and the diagrams are further well structured and complimentary to the text.

Very few edits and/or additions are suggested and these are included below:

- Introduction, Line 3: "However, it is an understudied..."
- Introduction, Paragraph 2: It would be helpful to the reader to have understand if there is information on *An. cydippis*. Is this known as a vector and does it have anthropophilic or zoophilic behaviours?
- Introduction, Paragraph 2: No mention is made of the unpublished claims of Green & Hunt that state there are 5 chromosomal forms of *An. squamosus* and also the publications that elude to multiple clades. This may be interesting to include to lend support to how understudied the species is.
- Data collection, Line 6: "...Eppendorf microcentrifuge tube (add company name)...." should be addressed.
- Results, paragraph 4: This paragraph requires some grammar editing such as "..that could be used for new primer design that would amply the ITS2 ...." should be corrected to "amplify".

Overall this is a very interesting and important article and is recommended for indexing.

Are the rationale for sequencing the genome and the species significance clearly described?

Yes

Are the protocols appropriate and is the work technically sound?

Yes

Are sufficient details of the sequencing and extraction, software used, and materials provided to allow replication by others?

Yes

Are the datasets clearly presented in a usable and accessible format, and the assembly and annotation available in an appropriate subject-specific repository?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Medical entomology, vector control, malaria, vector-borne disease

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.