



HAL
open science

Ribosomal RNA (rRNA) sequences from 33 globally distributed mosquito species for improved metagenomics and species identification

Cassandra Koh, Lionel Frangeul, Hervé Blanc, Carine Ngoagouni, Sébastien Boyer, Philippe Dussart, Nina Grau, Romain Girod, Jean-Bernard Duchemin, Maria-Carla Saleh

► To cite this version:

Cassandra Koh, Lionel Frangeul, Hervé Blanc, Carine Ngoagouni, Sébastien Boyer, et al.. Ribosomal RNA (rRNA) sequences from 33 globally distributed mosquito species for improved metagenomics and species identification. *eLife*, 2023, 12, pp.e82762. 10.7554/eLife.82762 . pasteur-03977506

HAL Id: pasteur-03977506

<https://pasteur.hal.science/pasteur-03977506v1>

Submitted on 7 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 Ribosomal RNA (rRNA) sequences from 33 globally distributed mosquito species for
2 improved metagenomics and species identification

3

4 Cassandra Koh^{1*}, Lionel Frangeul¹, Hervé Blanc¹, Carine Ngoagouni², Sébastien Boyer³, Philippe
5 Dussart⁴, Nina Grau⁵, Romain Girod⁵, Jean-Bernard Duchemin⁶ and Maria-Carla Saleh^{1*}

6 ¹ Viruses and RNA Interference Unit, Institut Pasteur, Université Paris Cité, CNRS UMR3569, Paris
7 F-75015, France

8 ² Medical Entomology Laboratory, Institut Pasteur de Bangui, Bangui PO Box 923, Central African
9 Republic

10 ³ Medical and Veterinary Entomology Unit, Institut Pasteur du Cambodge, Phnom Penh 12201,
11 Cambodia

12 ⁴ Virology Unit, Institut Pasteur du Cambodge, Phnom Penh 12201, Cambodia

13 ⁵ Medical Entomology Unit, Institut Pasteur de Madagascar, Antananarivo 101, Madagascar

14 ⁶ Vectopôle Amazonien Emile Abonnenc, Institut Pasteur de la Guyane, Cayenne 97306, French
15 Guiana

16 * To whom correspondence should be addressed.

17 Present address: Philippe Dussart, Institut Pasteur de Madagascar, Antananarivo 101, Madagascar;
18 Nina Grau, Sciences Economiques et Sociales de la Santé et Traitement de l'Information Médicale,
19 Faculté de Médecine, Marseille 13005, France

20 Corresponding authors: Cassandra Koh, cassandra.koh@pasteur.fr, +33 1 40 61 36 74; Maria-Carla
21 Saleh, carla.saleh@pasteur.fr, +33 1 45 68 85 47.

22 **ABSTRACT**

23 Total RNA sequencing (RNA-seq) is an important tool in the study of mosquitoes and the RNA viruses
24 they vector as it allows assessment of both host and viral RNA in specimens. However, there are two
25 main constraints. First, as with many other species, abundant mosquito ribosomal RNA (rRNA) serves
26 as the predominant template from which sequences are generated, meaning that the desired host and
27 viral templates are sequenced far less. Second, mosquito specimens captured in the field must be
28 correctly identified, in some cases to the sub-species level. Here, we generate mosquito ribosomal
29 RNA (rRNA) datasets which will substantially mitigate both of these problems. We describe a strategy
30 to assemble novel rRNA sequences from mosquito specimens and produce an unprecedented
31 dataset of 234 full-length 28S and 18S rRNA sequences of 33 medically important species from
32 countries with known histories of mosquito-borne virus circulation (Cambodia, the Central African
33 Republic, Madagascar, and French Guiana). These sequences will allow both physical and
34 computational removal of rRNA from specimens during RNAseq protocols. We also assess the utility
35 of rRNA sequences for molecular taxonomy and compare phylogenies constructed using rRNA
36 sequences versus those created using the gold standard for molecular species identification of
37 specimens—the mitochondrial *cytochrome c oxidase I* (COI) gene. We find that rRNA- and COI-
38 derived phylogenetic trees are incongruent and that 28S and concatenated 28S+18S rRNA
39 phylogenies reflect evolutionary relationships that are more aligned with contemporary mosquito
40 systematics. This significant expansion to the current rRNA reference library for mosquitoes will
41 improve mosquito RNA-seq metagenomics by permitting the optimization of species-specific rRNA
42 depletion protocols for a broader range of species and streamlining species identification by rRNA
43 sequence and phylogenetics.

44 **Keywords:** surveillance, RNA-seq, ribosomal RNA, molecular marker, metagenomics, mosquito

45 INTRODUCTION

46 Mosquitoes top the list of vectors for arthropod-borne diseases, being implicated in the transmission
47 of many human pathogens responsible for arboviral diseases, malaria, and lymphatic filariasis (WHO,
48 2017). Mosquito-borne viruses circulate in sylvatic (between wild animals) or urban (between
49 humans) transmission cycles driven by different mosquito species with their own distinct host
50 preferences. Although urban mosquito species are chiefly responsible for amplifying epidemics in
51 dense human populations, sylvatic mosquitoes maintain the transmission of these viruses among
52 forest-dwelling animal reservoir hosts and are implicated in spillover events when humans enter their
53 ecological niches (Valentine et al., 2019). Given that mosquito-borne virus emergence is preceded by
54 such spillover events, continuous surveillance and virus discovery in sylvatic mosquitoes is integral to
55 designing effective public health measures to pre-empt or respond to mosquito-borne viral epidemics.

56 Metagenomics on field specimens is the most powerful method in the toolkit for understanding
57 mosquito-borne disease ecology through the One Health lens (Webster et al., 2016). With next-
58 generation sequencing becoming more accessible, such studies have provided unprecedented
59 insights into the interfaces among mosquitoes, their environment, and their animal and human hosts.
60 As mosquito-associated viruses are mostly RNA viruses, RNA sequencing (RNA-seq) is especially
61 informative for surveillance and virus discovery. However, working with lesser studied mosquito
62 species poses several problems.

63 First, metagenomics studies based on RNA-seq are bedevilled by overabundant ribosomal RNAs
64 (rRNAs). These non-coding RNA molecules comprise at least 80% of the total cellular RNA
65 population (Gale & Crampton, 1989). Due to their length and their abundance, they are a sink for
66 precious next generation sequencing reads, decreasing the sensitivity of pathogen detection unless
67 depleted during library preparation. Yet the most common rRNA depletion protocols require prior
68 knowledge of rRNA sequences of the species of interest as they involve hybridizing antisense oligos
69 to the rRNA molecules prior to removal by ribonucleases (Fauver et al., 2019; Phelps et al., 2021) or
70 by bead capture (Kukutla et al., 2013). Presently, reference sequences for rRNAs are limited to only a
71 handful of species from three genera: *Aedes*, *Culex*, and *Anopheles* (Ruzzante et al., 2019). The lack
72 of reliable rRNA depletion methods could deter mosquito metagenomics studies from expanding their
73 sampling diversity, resulting in a gap in our knowledge of mosquito vector ecology. The inclusion of
74 lesser studied yet medically relevant sylvatic species is therefore imperative.

75 Second, species identification based on morphology is notoriously complicated for members of
76 species subgroups. This is especially the case among *Culex* subgroups. Sister species are often
77 sympatric and show at least some competence for a number of viruses, such as Japanese
78 encephalitis virus, St Louis encephalitic virus, and Usutu virus (Nchoutpouen et al., 2019). Although
79 they share many morphological traits, each of these species have distinct ecologies and host
80 preferences, thus the challenge of correctly identifying vector species can affect epidemiological risk
81 estimation for these diseases (Farajollahi et al., 2011). DNA molecular markers are often employed to
82 a limited degree of success to distinguish between sister species (Batovska et al., 2017; Zitra et al.,
83 2016).

84 To address the lack of full-length rRNA sequences in public databases, we sought to determine
85 the 28S and 18S rRNA sequences of a diverse set of Old and New World sylvatic mosquito species
86 from four countries representing three continents: Cambodia, the Central African Republic,
87 Madagascar, and French Guiana. These countries, due to their proximity to the equator, contain high
88 mosquito biodiversity (Foley et al., 2007) and have had long histories of mosquito-borne virus
89 circulation (Desdouits et al., 2015; Halstead, 2019; Héraud et al., 2022; Jacobi & Serie, 1972;
90 Ratsitorahina et al., 2008; Saluzzo et al., 2018; Zeller et al., 2016). Increased and continued
91 surveillance of local mosquito species could lead to valuable insights on mosquito virus biogeography.
92 Using a unique score-based read filtration strategy to remove interfering non-mosquito rRNA reads for
93 accurate *de novo* assembly, we produced a dataset of 234 novel full-length 28S and 18S rRNA
94 sequences from 33 mosquito species, 30 of which have never been recorded before.

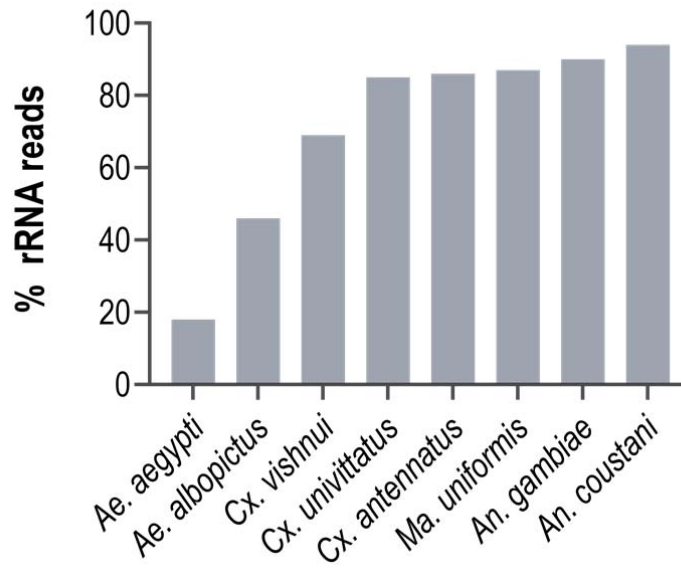
95 We also explored the functionality of 28S and 18S rRNA sequences as molecular markers by
96 comparing their performance to that of the mitochondrial *cytochrome c oxidase subunit I* (COI) gene
97 for molecular taxonomic and phylogenetic investigations. The COI gene is the most widely used DNA
98 marker for molecular species identification and forms the basis of the Barcode of Life Data System
99 (BOLD) (Hebert et al., 2003; Ratnasingham & Hebert, 2007). Presently, full-length rRNA sequences
100 are much less represented compared to other molecular markers. However, given the availability of
101 relevant reference sequences, 28S and concatenated 28S+18S rRNA sequences can be the better
102 approach for molecular taxonomy and phylogenetic studies. We hope that our sequence dataset, with
103 its species diversity and eco-geographical breadth, and the assembly strategy we describe would
104 further facilitate the use of rRNA as markers. In addition, this dataset enables the design of species-

105 specific oligos for cost-effective rRNA depletion for a broader range of mosquito species and
106 streamlined molecular species identification during RNA-seq.

107 **RESULTS**

108 **Poor rRNA depletion using a non-specific depletion method**

109 During library preparations of mosquito samples for RNA-seq, routinely used methods for depleting
110 rRNA are commercial kits optimised for human or mice samples (Belda et al., 2019; Bishop-Lilly et al.,
111 2010; Chandler et al., 2015; N. Kumar et al., 2012; Weedall et al., 2015; Zakrzewski et al., 2018) or
112 through 80–100 base pair antisense probe hybridisation followed by ribonuclease digestion (Fauver et
113 al., 2019; Phelps et al., 2021). In cases where the complete reference rRNA sequence of the target
114 species is not known, oligos would be designed based on the rRNA sequence of the closest related
115 species (25, this study). These methods should deplete the conserved regions of rRNA sequences.
116 However, the variable regions remain at abundances high enough to compromise RNA-seq output. In
117 our hands, we have found that using probes designed for the *Ae. aegypti* rRNA sequence followed by
118 RNase H digestion according to the protocol published by Morlan *et al.* (2012) produced poor
119 depletion in *Ae. albopictus*, and in Culicine and Anopheline species (Figure 1), in which between 46–
120 94% of reads post-depletion were ribosomal. Additionally, the lack of reference rRNA sequences
121 compromises the *in silico* clean-up of remaining rRNA reads from sequencing data, as reads
122 belonging to variable regions would not be removed. To solve this and to enable RNA-seq
123 metagenomics on a broader range of mosquito species, we performed RNA-seq to generate
124 reference rRNA sequences for 33 mosquito species representing 10 genera from Cambodia, the
125 Central African Republic, Madagascar, and French Guiana. Most of these species are associated with
126 vector activity for various pathogens in their respective ecologies (Table 1). In parallel, we sequenced
127 the mitochondrial COI gene to perform molecular species identification of our samples and to
128 comparatively evaluate the use of rRNA as a molecular marker (Figure 2).



129

130 **Figure 1. Percentage of rRNA reads in mosquito total RNA-seq data after depletion using**

131 **probes antisense to *Ae. aegypti* sequences.** Pools of 5 individual mosquitoes were ribodepleted by

132 probe hybridisation followed by RNase H digestion according to the protocol by Morlan *et al.* (2012).

133 Percentages of remaining rRNA reads were calculated from the number of rRNA reads over total

134 reads per sample pool. Depletion efficiency decreases with taxonomic distance from *Ae. aegypti*

135 underlining the need for reference sequences for species of interest.

136 **Table 1.** Mosquito species represented in this study and their vector status.

| Mosquito taxonomy* | Origin** | Collection site (ecosystem type) | Vector for*** | Reference |
|--|----------|----------------------------------|-----------------------------|--|
| <i>Aedes (Fredardsius) vittatus</i> | CF | rural (village) | ZIKV, CHIKV, YFV | (Diallo et al., 2020) |
| <i>Aedes (Ochlerotatus) scapularis</i> | GF | rural (village) | YFV | (Vasconcelos et al., 2001) |
| <i>Aedes (Ochlerotatus) serratus</i> | GF | rural (village) | YFV, OROV | (Cardoso et al., 2010; Romero-Alvarez & Escobar, 2018) |
| <i>Aedes (Stegomyia) aegypti</i> | CF | urban | DENV, ZIKV, CHIKV, YFV | (Kraemer et al., 2019) |
| <i>Aedes (Stegomyia) albopictus</i> | CF, KH | rural (village, nature reserve) | DENV, ZIKV, CHIKV, YFV, JEV | (Auerswald et al., 2021; Kraemer et al., 2019) |
| <i>Aedes (Stegomyia) simpsoni</i> | CF | rural (village) | YFV | (Mukwaya et al., 2000) |
| <i>Anopheles (Anopheles) baезai</i> | KH | rural (nature reserve) | unreported | – |
| <i>Anopheles (Anopheles) coustani</i> | MG, CF | rural (village) | RVFV, malaria | (Mwangangi et al., 2013; Nepomichene et al., 2018; Ratovonjato et al., 2011) |
| <i>Anopheles (Cellia) funestus</i> | MG, CF | rural (village) | ONNV, malaria | (Lutomiah et al., 2013; Tabue et al., 2017) |
| <i>Anopheles (Cellia) gambiae</i> | MG, CF | rural (village) | ONNV, malaria | (Brault et al., 2004) |
| <i>Anopheles (Cellia) squamosus</i> | MG | rural (village) | RVFV, malaria | (Ratovonjato et al., 2011; Stevenson et al., 2016) |

| | | | | |
|--|------------|---------------------------------|---|--|
| <i>Coquillettidia (Rhynchoetaenia) venezuelensis</i> | GF | rural (village) | OROV | (Travassos Da Rosa et al., 2017) |
| <i>Culex (Culex) antennatus</i> | MG | rural (village) | RVFV | (Nepomichene et al., 2018; Ratovonjato et al., 2011) |
| <i>Culex (Culex) duttoni</i> | CF | rural (village) | unreported | – |
| <i>Culex (Culex) neavei</i> | MG | rural (village) | USUV | (Nikolay et al., 2011) |
| <i>Culex (Culex) orientalis</i> | KH | rural (nature reserve) | JEV | (Kim et al., 2015) |
| <i>Culex (Culex) perexiguus</i> | MG | rural (village) | WNV, USUV | (Vázquez González et al., 2011) |
| <i>Culex (Culex) pseudovishnui</i> | KH | rural (nature reserve) | JEV | (Auerswald et al., 2021) |
| <i>Culex (Culex) quinquefasciatus</i> | MG, CF, KH | rural (village, nature reserve) | ZIKV, JEV, WNV, DENV, SLEV, RVFV, <i>Wuchereria bancrofti</i> | (Bhattacharya & Basu, 2016; Maquart et al., 2021; Ndiaye et al., 2016; Pereira Serra et al., 2016) |
| <i>Culex (Culex) tritaeniorhynchus</i> | MG, KH | rural (village, nature reserve) | JEV, WNV, RVFV | (Auerswald et al., 2021; Hayes et al., 1980; Jupp et al., 2002) |
| <i>Culex (Melanoconion) spissipes</i> | GF | rural (village) | VEEV | (Weaver et al., 2004) |
| <i>Culex (Melanoconion) portesi</i> | GF | rural (village) | VEEV, TONV | (Talaga et al., 2021; Weaver et al., 2004) |
| <i>Culex (Melanoconion) pedroi</i> | GF | rural (village) | EEEV, VEEV, MADV | (Talaga et al., 2021; M. J. Turell et al., 2008) |
| <i>Culex (Oculeomyia) bitaeniorhynchus</i> | MG, KH | rural (village, nature reserve) | JEV | (Auerswald et al., 2021) |
| <i>Culex (Oculeomyia) poicilipes</i> | MG | rural (village) | RVFV | (Ndiaye et al., 2016) |
| <i>Eretmapodites intermedius</i> | CF | rural (village) | unreported | – |
| <i>Limatus durhamii</i> | GF | rural (village) | ZIKV | (Barrio-Nuevo et al., 2020) |
| <i>Mansonia (Mansonia) titillans</i> | GF | rural (village) | VEEV, SLEV | (Hoyos-López et al., 2015; Michael J Turell, 1999) |
| <i>Mansonia (Mansonioides) indiana</i> | KH | rural (nature reserve) | JEV | (Arunachalam et al., 2004) |
| <i>Mansonia (Mansonioides) uniformis</i> | MG, CF, KH | rural (village, nature reserve) | RVFV, <i>Wuchereria bancrofti</i> | (Lutumiah et al., 2013; Ughasi et al., 2012) |
| <i>Mimomyia (Etorleptomyia) mediolineata</i> | MG | rural (village) | unreported | – |
| <i>Psorophora (Janthinosoma) ferox</i> | GF | rural (village) | ROCV | (Mitchell et al., 1986) |
| <i>Uranotaenia (Uranotaenia) geometrica</i> | GF | rural (village) | unreported | – |

137 * () indicates subgenus

138 ** Origin countries are listed as their ISO alpha-2 codes: Central African Republic, CF; Cambodia, KH;

139 Madagascar, MG; French Guiana, GF.

140 ** dengue virus, DENV; Zika virus, ZIKV; chikungunya virus, CHIKV; Yellow Fever virus, YFV; Oropouche virus,

141 OROV; Japanese encephalitis virus, JEV; Rift Valley Fever virus, RVFV; O’Nyong Nyong virus, ONNV; Usutu

142 virus, USUV; West Nile virus, WNV; Saint Louis encephalitis virus, SLEV; Venezuelan equine encephalitis

143 virus, VEEV; Tonate virus, TONV; Eastern equine encephalitis virus, EEEV; Madariaga virus, MADV; Rocio

144 virus, ROCV.

145 **rRNA reads filtering and sequence assembly**

146 Assembling Illumina reads to reconstruct rRNA sequences from total mosquito RNA is not a

147 straightforward task. Apart from host rRNA, total RNA samples also contain rRNA from other

148 organisms associated with the host (microbiota, external parasites, or ingested diet). As rRNA
149 sequences share high homology in conserved regions, Illumina reads (150 bp) from non-host rRNA
150 can interfere with the contig assembly of host 28S and 18S rRNA.

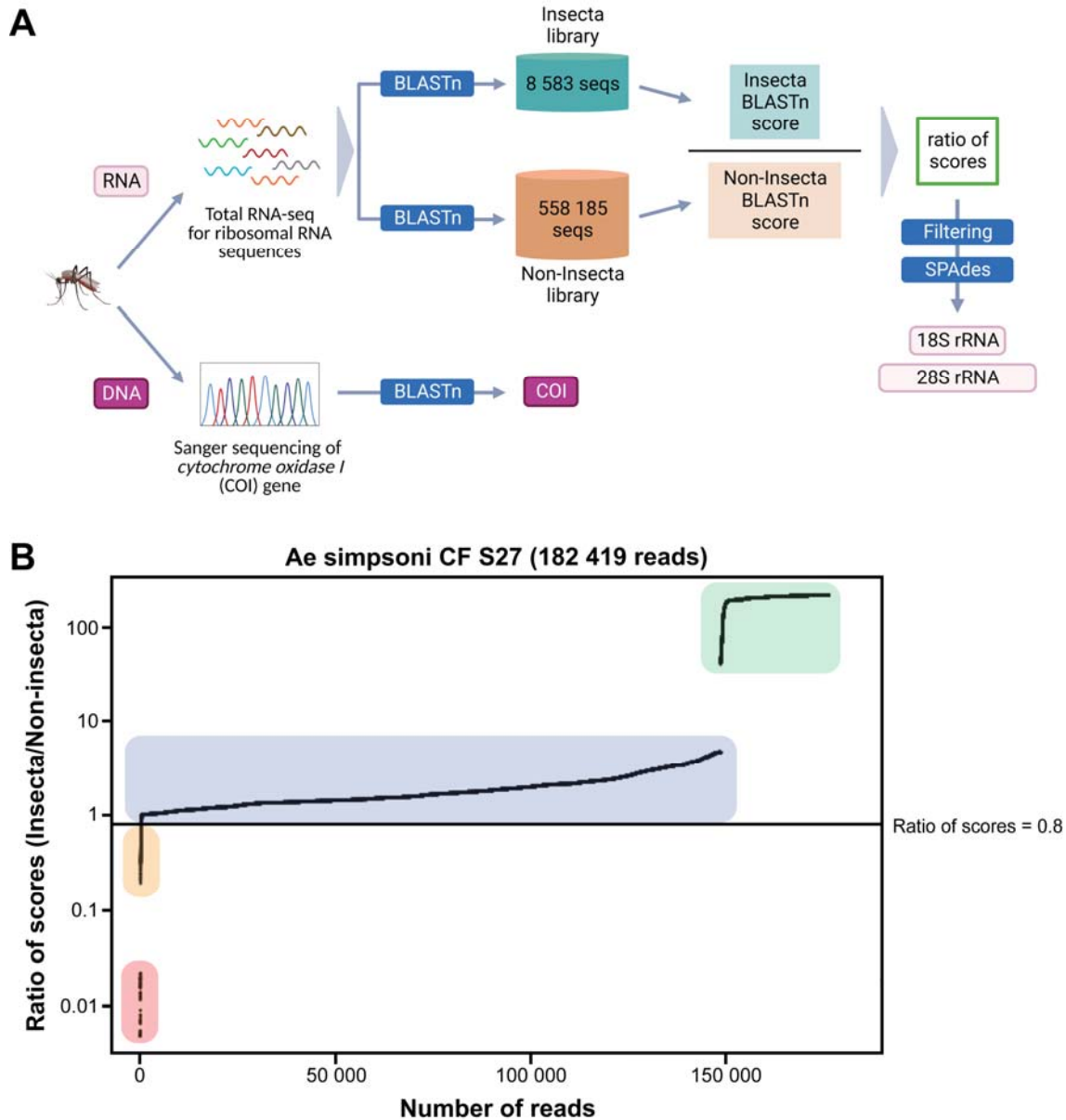
151 Our score-based filtration strategy, described in detail in the Methods section, allowed us to
152 bioinformatically remove interfering rRNA reads and achieve successful *de novo* assembly of 28S and
153 18S rRNA sequences for all our specimens. Briefly, for each Illumina read, we computed a ratio of
154 BLAST scores against an Insecta library over scores against a Non-Insecta library (Figure 2A). Based
155 on their ratio of scores, reads could be segregated into four categories (Figure 2B): (i) reads mapping
156 only to the Insecta library, (ii) reads mapping better to the Insecta relative to Non-Insecta library, (iii)
157 reads mapping better to the Non-Insecta relative to the Insecta library, and (iv) reads mapping only to
158 the Non-Insecta library. By applying a conservative threshold at 0.8 to account for the non-
159 exhaustiveness of the SILVA database, we removed reads that likely do not originate from mosquito
160 rRNA. Notably, 15 of our specimens were engorged with vertebrate blood, a rich source of non-
161 mosquito rRNA (Appendix 1—table 1). The successful assembly of complete 28S and 18S rRNA
162 sequences for these specimens demonstrates that this strategy performs as expected even with high
163 amounts of non-host rRNA reads. This is particularly important in studies on field-captured
164 mosquitoes as females are often sampled already having imbibed a blood meal or captured using the
165 human landing catch technique.

166 We encountered challenges for three specimens morphologically identified as *Ma. africana*
167 (Specimen ID S33–S35) (Appendix 1—table 1). COI amplification by PCR did not produce any
168 product, hence COI sequencing could not be used to confirm species identity. In addition, the genome
169 assembler SPAdes (Bankevich et al., 2012) was only able to assemble partial length rRNA contigs,
170 despite the high number of reads with high scores against the Insecta library. Among other *Mansonia*
171 specimens, the partial length contigs shared the highest similarity with contigs obtained from sample
172 “Ma uniformis CF S51”. We then performed a guided assembly using the 28S and 18S sequences of
173 this specimen as references, which successfully produced full-length contigs. In two of these
174 specimens (Specimen ID S34 and S35), our assembly initially produced two sets of 28S and 18S
175 rRNA sequences, one of which was similar to mosquito rRNA with low coverage and another with ten-
176 fold higher coverage and 95% nucleotide sequence similarity to a water mite of genus *Horreolanus*
177 known to parasitize mosquitoes. Our success in obtaining rRNA sequences for mosquito and water

178 mite shows that our strategy can be applied to metabarcoding studies where the input material
179 comprises multiple insect species, provided that appropriate reference sequences of the target
180 species or of a close relative are available.

181 Altogether, we were able to assemble 122 28S and 114 18S full-length rRNA sequences for 33
182 mosquito species representing 10 genera sampled from four countries across three continents. This
183 dataset contains, to our knowledge, the first records for 30 mosquito species and for seven genera:
184 *Coquillettidia*, *Mansonia*, *Limatus*, *Mimomyia*, *Uranotaenia*, *Psorophora*, and *Eretmapodites*.

185 Individual GenBank accession numbers for these sequences and specimen information are listed in
186 Appendix 1—table 1.



187

188 **Figure 2. Novel mosquito rRNA sequences were obtained using a unique filtering method. (A)**

189 Schematic of sequencing and bioinformatics analysis performed in this study to obtain full-length 18S
 190 and 28S rRNA sequences as well as COI DNA sequences. Nucleic acids were isolated from mosquito
 191 specimens for next generation (for rRNA) or Sanger (for COI) sequencing. Two in-house libraries
 192 were created from the SILVA rRNA gene database: Insecta and Non-Insecta, which comprises 8 585
 193 sequences and 558 185 sequences, respectively. Following BLASTn analysis against these two
 194 libraries, each RNA-seq read is assigned a ratio of BLASTn scores to describe their relative
 195 nucleotide similarity to insect rRNA sequences. Based on these ratios of scores, RNA-seq reads can

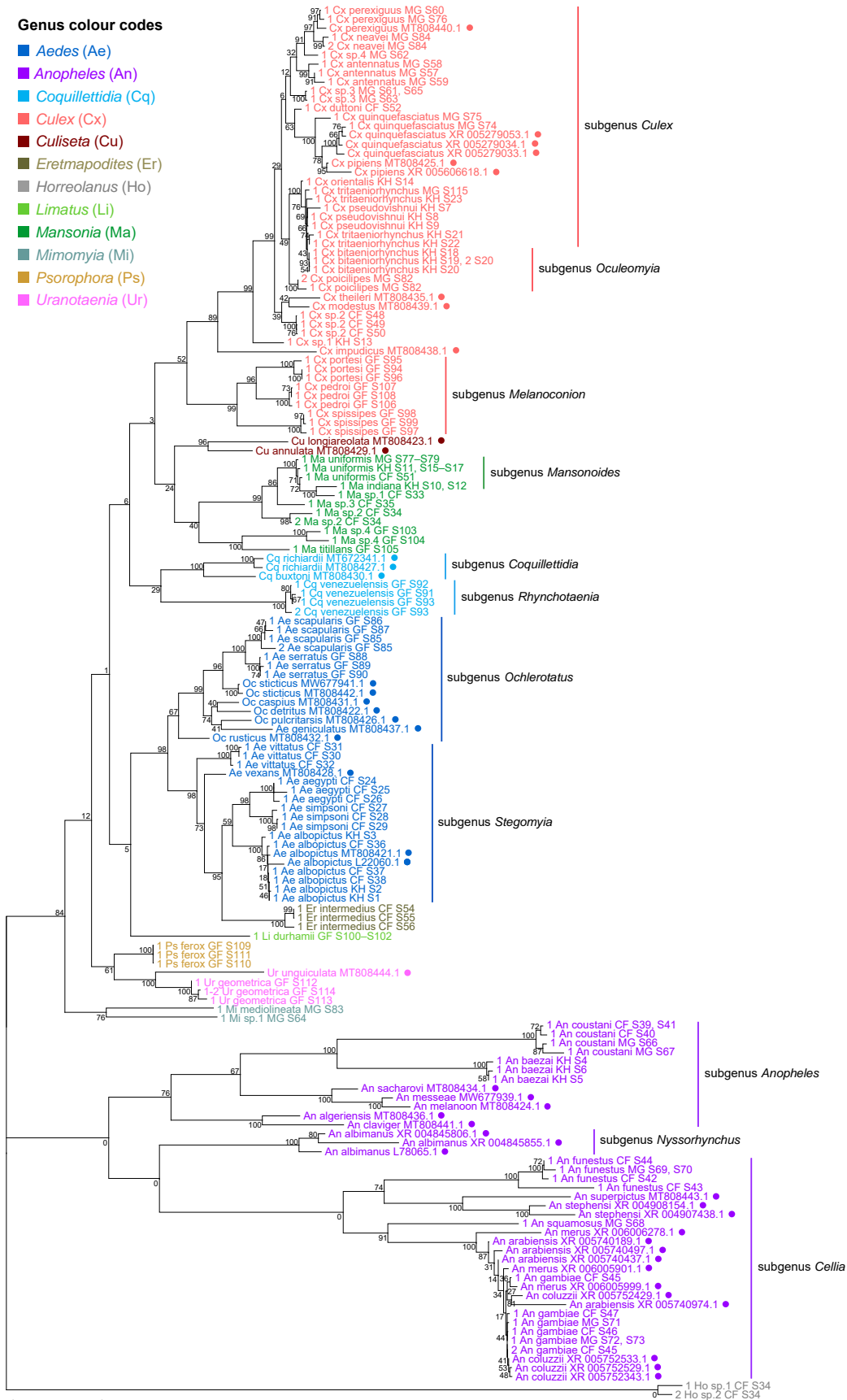
196 then be filtered to remove non-mosquito reads prior to assembly with SPAdes to give full-length 18S
197 and 28S rRNA sequences. Image created with Biorender.com. **(B)** Based on their ratio of scores,
198 reads can be segregated into four categories, as shown on this ratio of scores vs. number of reads
199 plot for the representative specimen “Ae simpsoni CF S27”: (i) reads with hits only in the Insecta
200 library (shaded in green), (ii) reads with a higher score against the Insecta library (shaded in blue), (iii)
201 reads with a higher score against the Non-Insecta library (shaded in yellow), and (iv) reads with no
202 hits in the Insecta library (shaded in red). We applied a conservative threshold at 0.8, indicated by the
203 black horizontal line, where only reads above this threshold are used in the assembly with SPAdes.
204 For this given specimen, 175 671 reads (96.3% of total reads) passed the ≥ 0.8 cut-off, 325 reads
205 (0.18% of total reads) had ratios of scores < 0.8 , while 6 423 reads (3.52%) did not have hits against
206 the Insecta library.

207 **Comparative phylogeny of novel rRNA sequences relative to existing records**

208 To verify the assembly accuracy of our rRNA sequences, we constructed a comprehensive
209 phylogenetic tree from the full-length 28S rRNA sequences generated from our study and included
210 relevant rRNA sequences publicly available from GenBank (Figure 3). We applied a search criterion
211 for GenBank sequences with at least 95% coverage of our sequence lengths (~4000 bp), aiming to
212 represent as many species or genera as possible. Although we rarely found records for the same
213 species included in our study, the resulting tree showed that our 28S sequences generally clustered
214 according to their respective species and subgenera, supported by moderate to good bootstrap
215 support at terminal nodes. Species taxa generally formed monophyletic clades, with the exception of
216 *An. gambiae* and *Cx. quinquefasciatus*. *An. gambiae* 28S rRNA sequences formed a clade with
217 closely related sequences from *An. arabiensis*, *An. merus*, and *An. coluzzii*, suggesting unusually
218 high interspecies homology for Anophelines or other members of subgenus *Cellia* (Figure 3, in purple,
219 subgenus *Cellia*). Meanwhile, *Cx. quinquefasciatus* 28S rRNA sequences formed a taxon
220 paraphyletic to sister species *Cx. pipiens* (Figure 3, in coral, subgenus *Culex*).

Genus colour codes

- *Aedes* (Ae)
- *Anopheles* (An)
- *Coquillettia* (Cq)
- *Culex* (Cx)
- *Culiseta* (Cu)
- *Eretmapodites* (Er)
- *Horreolanus* (Ho)
- *Limatus* (Li)
- *Mansonia* (Ma)
- *Mimomyia* (Mi)
- *Psorophora* (Ps)
- *Uranotaenia* (Ur)



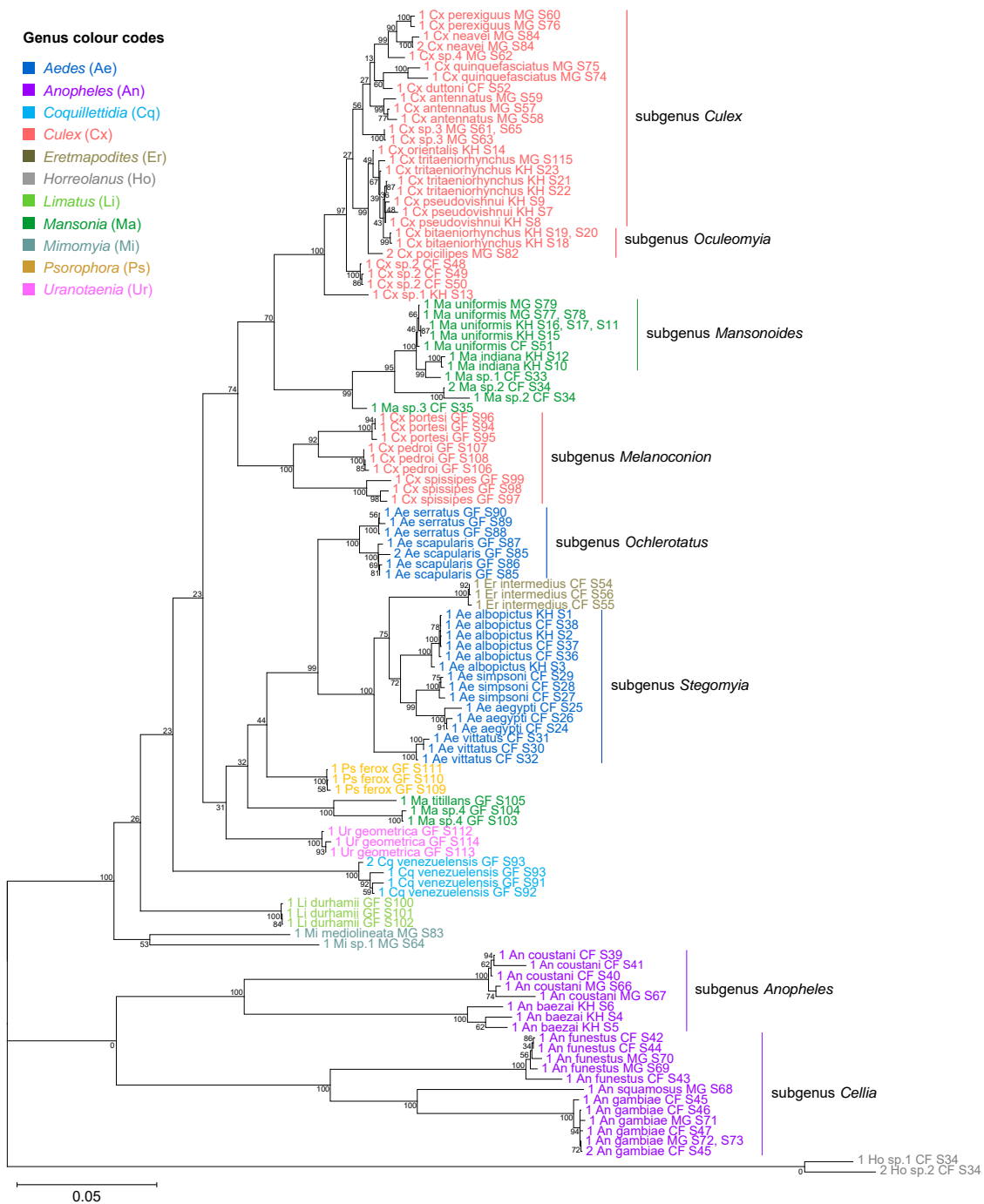
222 **Figure 3. 28S sequences generated from this study clustered with conspecifics or congenics**
223 **from existing GenBank records.** A rooted phylogenetic tree based on full-length 28S sequences
224 (3900 bp) from this study and from GenBank was inferred using the maximum-likelihood method and
225 constructed to scale in MEGA X (S. Kumar et al., 2018) using an unknown *Horreolanus* species found
226 among our samples as an outgroup. Values at each node indicate bootstrap support (%) from 500
227 replications. Sequences from GenBank are annotated with filled circles and their accession numbers
228 are shown. For sequences from this study, each specimen label contains information on taxonomy,
229 origin (in 2-letter country codes), and specimen ID number. Some specimens produced up to two
230 consensus 28S sequences; this is indicated by the numbers 1 or 2 in the beginning of the specimen
231 label. Specimen genera are indicated by colour: *Culex* in coral, *Anopheles* in purple, *Aedes* in dark
232 blue, *Mansonia* in dark green, *Culiseta* in maroon, *Limatus* in light green, *Coquillettidia* in light blue,
233 *Psorophora* in yellow, *Mimomyia* in teal, *Uranotaenia* in pink and *Eretmapodites* in brown. Scale bar
234 at 0.05 is shown.

235 28S rRNA sequence-based phylogenetic reconstructions (Figure 3, with GenBank sequences;
236 Figure 4—figure supplement 1, this study only) showed marked incongruence to that of 18S rRNA
237 sequences (Figure 4—figure supplement 2). Although all rRNA trees show the bifurcation of family
238 *Culicidae* into subfamilies *Anophelinae* (genus *Anopheles*, in purple) and *Culicinae* (all other genera),
239 the recovered intergeneric phylogenetic relationships vary between the 28S and 18S rRNA trees and
240 are weakly supported. The 18S rRNA tree also exhibited several taxonomic anomalies: (i) the lack of
241 definitive clustering by species within the *Culex* subgenus (in coral) (ii) the lack of distinction between
242 18S rRNA sequences of *Cx. pseudovishnui* and *Cx. tritaeniorhynchus* (in coral); (iii) the placement of
243 Ma sp. 3 CF S35 (in dark green) within a *Culex* clade; and (iv) the lack of a monophyletic *Mimomyia*
244 clade (in teal) (Figure 4—figure supplement 2). However, 28S and 18S rRNA sequences are encoded
245 by linked loci in rDNA clusters and should not be analysed separately.

246 Indeed, when concatenated 28S+18S rRNA sequences were generated from the same specimens
247 (Figure 4), the phylogenetic tree resulting from these sequences more closely resembles the 28S tree
248 (Figure 3) with regard to the basal position of the *Mimomyia* clade (in teal) within the *Culicinae*
249 subfamily with good bootstrap support in either tree (84% in 28S rRNA tree, 100% in concatenated
250 28S+18S rRNA tree). For internal nodes, bootstrap support values were higher in the concatenated
251 tree compared to the 28S tree. Interestingly, the 28S+18S rRNA tree formed an *Aedini* tribe-clade

252 encompassing taxa from genera *Psorophora* (in yellow), *Aedes* (in dark blue), and *Eretmapodites* (in
253 brown), possibly driven by the inclusion of 18S rRNA sequences. Concatenation also resolved the
254 anomalies found in the 18S rRNA tree and added clarity to the close relationship between *Culex* (in
255 coral) and *Mansonia* (in dark green) taxa. Of note, relative to the 28S tree (Figure 3) the *Culex* and
256 *Mansonia* genera are no longer monophyletic in the concatenated 28S+18S rRNA tree (Figure 4).
257 Genus *Culex* is paraphyletic with respect to subgenus *Mansonoides* of genus *Mansonia* (Figure 3).
258 *Ma. titillans* and *Ma. sp. 4*, which we suspect to be *Ma. pseudotitillans*, always formed a distinct branch
259 in 28S or 18S rRNA phylogenies, thus possibly representing a clade of subgenus *Mansonia*.

260 The concatenated 28S+18S rRNA tree (Figure 4) recapitulates what is classically known about the
261 systematics of our specimens, namely (i) the early divergence of subfamily *Anophelinae* from
262 subfamily *Culicinae*, (ii) the division of genus *Anopheles* (in purple) into two subgenera, *Anopheles*
263 and *Cellia*, (iii) the division of genus *Aedes* (in dark blue) into subgenera *Stegomyia* and *Ochlerotatus*,
264 (iv) the divergence of the monophyletic subgenus *Melanoconion* within the *Culex* genus (in coral)
265 (Harbach, 2007; Harbach & Kitching, 2016).



266

267 **Figure 4. Concatenating 28S and 18S rRNA sequences produces phylogenetic relationships**
 268 **that are concordant with classical *Culicidae* systematics with higher bootstrap support than**
 269 **28S sequences alone.** This phylogenetic tree based on concatenated 28S+18S rRNA sequences
 270 (3900+1900 bp) generated from this study was inferred using the maximum-likelihood method and
 271 constructed to scale using MEGA X (S. Kumar et al., 2018) using an unknown *Horreolanus* species
 272 found among our samples as an outgroup. Values at each node indicate bootstrap support (%) from

273 500 replications. Each specimen label contains information on taxonomy, origin (as indicated in 2-
274 letter country codes), and specimen ID number. Some specimens produced up to two consensus
275 28S+18S rRNA sequences; this is indicated by the numbers 1 or 2 in the beginning of the specimen
276 label. Specimen genera are indicated by colour: *Culex* in coral, *Anopheles* in purple, *Aedes* in dark
277 blue, *Mansonia* in dark green, *Limatus* in light green, *Coquillettidia* in light blue, *Psorophora* in yellow,
278 *Mimomyia* in teal, *Uranotaenia* in pink and *Eretmapodites* in brown. Scale bar at 0.05 is shown.

279 **rRNA as a molecular marker for taxonomy and phylogeny**

280 We sequenced a 621 bp region of the COI gene to confirm morphological species identification of our
281 specimens and to compare the functionality of rRNA and COI sequences as molecular markers for
282 taxonomic and phylogenetic investigations. COI sequences were able to unequivocally determine the
283 species identity in most specimens except for the following cases. *An. coustani* COI sequences from
284 our study regardless of specimen origin shared remarkably high nucleotide similarity (>98%) with
285 several other *Anopheles* species such as *An. rhodesiensis*, *An. rufipes*, *An. ziemanni*, *An. tenebrosus*,
286 although *An. coustani* remained the most frequent and closest match. In the case of *Ae. simpsoni*,
287 three specimens had been morphologically identified as *Ae. opok* although their COI sequences
288 showed 97–100% similarity to that of *Ae. simpsoni*. As GenBank held no records of *Ae. opok* COI at
289 the time of this study, we instead aligned the putative *Ae. simpsoni* COI sequences against two sister
290 species of *Ae. opok*: *Ae. luteocephalus* and *Ae. africanus*. We found they shared only 90% and 89%
291 similarity, respectively. Given this significant divergence, we concluded these specimens to be *Ae.*
292 *simpsoni*. Ambiguous results were especially frequent among *Culex* specimens belonging to the *Cx.*
293 *pipiens* or *Cx. vishnui* subgroups, where the query sequence differed with either of the top two hits by
294 a single nucleotide. For example, between *Cx. quinquefasciatus* and *Cx. pipiens* of the *Cx. pipiens*
295 subgroup, and between *Cx. vishnui* and *Cx. tritaeniorhynchus* of the *Cx. vishnui* subgroup.

296 Among our three specimens of *Ma. titillans*, two appeared to belong to a single species that is
297 different from but closely related to *Ma. titillans*. We surmised that these specimens could instead be
298 *Ma. pseudotitillans* based on morphological similarity but were not able to verify this by molecular
299 means as no COI reference sequence is available for this species. These specimens are hence
300 putatively labelled as “Ma sp.4”.

301 Phylogenetic reconstruction based on the COI sequences showed clustering of all species taxa
302 into distinct clades, underlining the utility of the COI gene in molecular taxonomy (Figure 5)(Hebert et

303 al., 2003; Ratnasingham & Hebert, 2007). However, species delineation among members of *Culex*
304 subgroups were not as clear cut, although sister species were correctly placed as sister taxa (Figure
305 5, in coral). This is comparable to the 28S+18S rRNA tree (Figure 4, in coral) and is indicative of lower
306 intraspecies distances relative to interspecies distances.

307 To evaluate the utility of 28S and 18S rRNA sequences for molecular taxonomy, we used the
308 28S+18S rRNA tree to discern the identity of six specimens for which COI sequencing could not be
309 performed. These specimens include three unknown *Mansonia* species (Specimen ID S33–S35), a
310 *Ma. uniformis* (Specimen ID S51), an *An. gambiae* (Specimen ID S47), and a *Ur. geometrica*
311 (Specimen ID S113) (Appendix 1—table 1). Their positions in the 28S+18S rRNA tree relative to
312 adjacent taxa confirms the morphological identification of all six specimens to the genus level and, for
313 three of them, to the species level (Figure 4; *Mansonia* in dark green, *Anopheles* in purple,
314 *Uranotaenia* in pink).

315 The phylogenetic relationships indicated by the COI tree compared to the 28S+18S rRNA tree
316 present only few points of similarity, with key differences summarised in Table 2. COI-based
317 phylogenetic inference indeed showed clustering of generic taxa into monophyletic clades albeit with
318 very weak bootstrap support, except for genera *Culex* and *Mansonia* (Figure 5; *Culex* in coral,
319 *Mansonia* in dark green). Contrary to the 28S+18S rRNA tree (Figure 4), *Culex* subgenus
320 *Melanoconion* was depicted as a polyphyletic taxon with *Cx. spissipes* being a part of the greater
321 *Culicini* clade with members from subgenera *Oculeomyia* and *Culex* while *Cx. pedroi* and *Cx. portesi*
322 formed a distantly related clade. Among the *Mansonia* specimens, the two unknown *Ma* sp.4
323 specimens were not positioned as the nearest neighbours of *Ma. titillans* and instead appeared to
324 have diverged earlier from most of the other taxa from the *Culicidae* family. Notably, the COI
325 sequences of genus *Anopheles* (Figure 5, in purple) is not basal to the other members of *Culicidae*
326 and is instead shown to be sister to *Culex* COI sequences (8% bootstrap support). This is a direct
327 contrast to what is suggested by the rRNA phylogenies (Figures 3 and 4, Figure 4—figure
328 supplements 1 and 2; *Anopheles* in purple), which suggests *Culex* (in coral) rRNA sequences to be
329 among the most recently diverged. Bootstrap support for the more internal nodes of the COI trees
330 were remarkably low compared to those of rRNA-based trees.

331 In all rRNA trees, it is clear that the interspecific and intersubgeneric evolutionary distances within
332 the genus *Anopheles* are high relative to any other genera, indicating a greater degree of divergence

333 (Figure 3, Figure 3—figure supplement 1, Figure 4, Figure 4—figure supplements 1 and 2; *Anopheles*
 334 in purple). This is evidenced by the longer branch lengths connecting Anopheline species-clades to
 335 the node of the most recent common ancestor for subgenera *Anopheles* and *Cellia*. This feature is not
 336 evident in the COI tree, where the Anopheline interspecies distances are comparable to those within
 337 the *Culex*, *Aedes*, and *Mansonia* taxa (Figure 5; *Anopheles* in purple, *Culex* in coral, *Aedes* in dark
 338 blue, *Mansonia* in dark green).



339
 340 **Figure 5. COI sequences cluster by species but show phylogenetic relationships that contrast**
 341 **those derived from rRNA trees. A phylogenetic tree based on COI sequences (621–699 bp) was**

342 inferred using the maximum-likelihood method and constructed to scale using MEGA X (S. Kumar et
 343 al., 2018) with three water mite species to serve as outgroups. Outgroup sequences obtained from
 344 GenBank are annotated with filled circles and their accession numbers are shown. Values at each
 345 node indicate bootstrap support (%) from 500 replications. Each specimen label contains information
 346 on taxonomy, origin (as indicated in 2-letter country codes), and specimen ID. Specimen genera are
 347 indicated by colour: *Culex* in coral, *Anopheles* in purple, *Aedes* in dark blue, *Mansonia* in dark green,
 348 *Limatus* in light green, *Coquillettidia* in light blue, *Psorophora* in yellow, *Mimomyia* in teal, *Uranotaenia*
 349 in pink and *Eretmapodites* in brown. Scale bar at 0.05 is shown.

350 **Table 2.** Summary of differences between rRNA and COI phylogenies.

| Taxa | 28S+18S rRNA phylogeny (Figure 4) | COI phylogeny (Figure 5) |
|-----------------------------------|--|--|
| The <i>Anopheles</i> genus | forms a clade that is basal to the all other members of family <i>Culicidae</i> ; interspecies branch lengths are notably long | forms a sister clade to the <i>Culex</i> genus, and is depicted to have diverged more recently; interspecies branch lengths are comparable to that of other genera |
| The <i>Ur. geometrica</i> species | forms a clade within the <i>Culicinae</i> subfamily lineage | forms a clade that is basal to the all other members of family <i>Culicidae</i> |
| The <i>Aedini</i> tribe | forms a monophyletic clade comprising the genera <i>Aedes</i> , <i>Eretmapodites</i> , and <i>Psorophora</i> , with the latter being an early divergent lineage | does not form a monophyletic clade; the <i>Psorophora</i> clade is placed among <i>Aedes</i> taxa and the <i>Eretmapodites</i> clade is sister to a <i>Culex</i> subgenus <i>Melanoconion</i> clade |
| The <i>Culex</i> genus | splits into two monophyletic clades with the three French Guyanese species forming a closely-related minor clade | splits into two clades with two out of three French Guyanese species (<i>Cx. pedroi</i> and <i>Cx. portesi</i>) forming a distantly-related minor clade, while the third (<i>Cx. spissipes</i>) is a part of the greater clade |
| The <i>Mansonia</i> genus | is a polyphyletic group comprising two clades with the two French Guyanese taxa forming a distantly-related minor clade; the major clade is placed among <i>Culex</i> taxa | forms a subgenus <i>Mansonoidea</i> clade as per the 28S+18S rRNA tree but the French Guyanese taxa do not cluster together; is depicted to have diverged earlier relative to other taxa in the assemblage |
| The Ma sp.4 species | forms a sister clade to <i>Ma. titillans</i> as part of a minor French Guyanese <i>Mansonia</i> clade | does not form a sister clade to <i>Ma. titillans</i> ; instead is shown to have diverged earlier than all other members of family <i>Culicidae</i> after <i>Ur. geometrica</i> |

351

352 **On *Culex* subgroups**

353 *Culex* (subgenus *Culex*) specimens of this study comprise several closely related sister species
354 belonging to the *Cx. vishnui* and *Cx. univittatus* subgroups, which are notoriously difficult to
355 differentiate based on morphology. Accordingly, in the 28S+18S rRNA (Figure 4, in coral) and COI
356 (Figure 5, in coral) trees these species and their known sister species were clustered together within
357 the *Culex* (subgenus *Culex*) clade: *Cx. tritaeniorhynchus* with *Cx. pseudovishnui* (*Cx. vishnui*
358 subgroup); *Cx. perexiguus* with *Cx. neavei* (*Cx. univittatus* subgroup).

359 The use of the COI sequence to distinguish between members of the *Culex* subgroups was
360 limited. For example, for the two *Cx. quinquefasciatus* samples in our taxonomic assemblage
361 (Specimen ID S74 and S75) (Appendix 1—table 1), BLAST analyses of their COI sequences revealed
362 they are a single nucleotide away from *Cx. pipiens* or *Cx. quinquefasciatus* COI sequences (Appendix
363 2—table 1). In the 28S rRNA tree with GenBank sequences (Figure 3), two *Cx. pipiens* GenBank
364 sequences formed a clade sister to another containing three *Cx. quinquefasciatus* GenBank
365 sequences and the “*Cx quinquefasciatus* MG S74” sequence with 78% bootstrap support. This is in
366 accordance with other studies examining mitochondrial sequences (Sun et al., 2019) and
367 morphological attributes (Harbach et al., 2017). This shows that the 28S rRNA sequence can
368 distinguish the two species and confirms that “*Cx quinquefasciatus* MG S74” is indeed a *Cx.*
369 *quinquefasciatus* specimen. However, “*Cx quinquefasciatus* MG S75” is shown to be basal from other
370 sequences within this *Cx. pipiens* subgroup-clade with 100% bootstrap support. Given that *Cx.*
371 *quinquefasciatus* and *Cx. pipiens* are known to interbreed, it is plausible that this individual is a hybrid
372 of the two species (Farajollahi et al., 2011).

373 **DISCUSSION**

374 RNA-seq metagenomics on field-captured sylvatic mosquitoes is a valuable tool for tracking
375 mosquito viruses through surveillance and virus discovery. However, the lack of reference rRNA
376 sequences hinders good oligo-based depletion and efficient clean-up of RNA-seq data. Additionally,
377 *de novo* assembly of rRNA sequences is complicated due to regions that are highly conserved across
378 all distantly related organisms that could be present in a single specimen, i.e., microbiota, parasites,
379 or vertebrate blood meal. Hence, we sought to establish a method to bioinformatically filter out non-
380 host rRNA reads for the accurate assembly of novel 28S and 18S rRNA reference sequences.

381 We found that phylogenetic reconstructions based on 28S sequences or concatenated 28S+18S
382 rRNA sequences were able to correctly cluster mosquito taxa according to species and corroborate
383 current mosquito classification. This demonstrates that our bioinformatics methodology reliably
384 generates bona fide 28S and 18S rRNA sequences, even in specimens parasitized by water mites or
385 engorged with vertebrate blood. Further, we were able to use 28S+18S rRNA taxonomy for molecular
386 species identification when COI sequences were unavailable or ambiguous, thus supporting the use
387 of rRNA sequences as a marker. They have the advantage of circumventing the need to additionally
388 isolate and sequence DNA from specimens, as RNA-seq reads can be directly mapped against
389 reference sequences. Post-depletion, in our hands there are sufficient numbers of remaining reads
390 (5–10% of reads per sample) for assembly of complete rRNA contigs (unpublished data).

391 Phylogenetic inferences based on 28S or 18S rRNA sequences alone do not recover the same
392 interspecific relationships (Figure 4—figure supplements 1 and 2). Relative to 28S sequences, we
393 observed more instances where multiple specimens have near-identical 18S rRNA sequences. This
394 can occur for specimens belonging to the same species, but also for conspecifics sampled from
395 different geographic locations, such as *An. coustani*, *An. gambiae*, or *Ae. albopictus*. More rarely,
396 specimens from the same species subgroup, such as *Cx. pseudovishnui* and *Cx. tritaeniorhynchus*,
397 also shared 18S rRNA sequences. This was surprising given that the 18S rRNA sequences in our
398 dataset is 1900 bp long. Concatenation of 28S and 18S rRNA sequences resolved this issue,
399 enabling species delineation even among sister species of *Culex* subgroups, where morphological
400 identification meets its limits.

401 In Cambodia and other parts of Asia, the *Cx. vishnui* subgroup includes *Cx. tritaeniorhynchus*, *Cx.*
402 *vishnui*, and *Cx. pseudovishnui*, which are important vectors of JEV (Maquart & Boyer, 2022). The
403 former two were morphologically identified in our study but later revealed by COI sequencing to be a
404 sister species. Discerning sister species of the *Cx. pipiens* subgroup is further complicated by
405 interspecific breeding, with some populations showing genetic introgression to varying extents (Cornel
406 et al., 2003). The seven sister species of this subgroup are practically indistinguishable based on
407 morphology and require molecular methods to discern (Farajollahi et al., 2011; Zittra et al., 2016).
408 Indeed, the 621 bp COI sequence amplified in our study did not contain enough nucleotide
409 divergence to allow clear identification, given that the COI sequence of *Cx. quinquefasciatus*
410 specimens differed from that of *Cx. pipiens* by a single nucleotide. Batovska et al. (2017) found that

411 even the Internal Transcribed Spacer 2 (ITS2) rDNA region, another common molecular marker,
412 could not differentiate the two species. Other DNA molecular markers such as nuclear *Ace-2* or *CQ11*
413 genes (Aspen & Savage, 2003; Zittra et al., 2016) or *Wolbachia pipientis* infection status (Cornel et
414 al., 2003) are typically employed in tandem. In our study, 28S rRNA sequence-based phylogeny
415 validated the identity of specimen “*Cx quinquefasciatus* MG S74” (Figure 3, in coral) and suggested
416 that specimen “*Cx quinquefasciatus* MG S75” might have been a *pipiens-quinquefasciatus* hybrid.
417 These examples demonstrate how 28S rRNA sequences, concatenated with 18S rRNA sequences or
418 alone, contain enough resolution to differentiate between *Cx. pipiens* and *Cx. quinquefasciatus*.
419 rRNA-based phylogeny thus allows for more accurate species identification and ecological
420 observations in the context of disease transmission. Additionally, tracing the genetic flow across
421 hybrid populations within the *Cx. pipiens* subgroup can inform estimates of vectorial capacity for each
422 species. As only one or two members from the *Cx. pipiens* and *Cx. vishnui* subgroups were
423 represented in our taxonomic assemblage, an explicit investigation including all member species of
424 these subgroups in greater sample numbers is warranted to further test the degree of accuracy with
425 which 28S and 18S rRNA sequences can delineate sister species.

426 Our study included French Guianese *Culex* species *Cx. spissipes* (group *Spissipes*), *Cx. pedroi*
427 (group *Pedroi*), and *Cx. portesi* (group *Vomerifer*). These species belong to the New World subgenus
428 *Melanoconion*, section *Spissipes*, with well-documented distribution in North and South Americas
429 (Sirivanakarn, 1982) and are vectors of encephalitic alphaviruses EEEV and VEEV among others
430 (Talaga et al., 2021; M. J. Turell et al., 2008; Weaver et al., 2004). Indeed, our rooted rRNA and COI
431 trees showed the divergence of the three *Melanoconion* species from the major *Culex* clade
432 comprising species broadly found across Africa and Asia (Auerswald et al., 2021; Farajollahi et al.,
433 2011; Nchoutpouen et al., 2019; Takhampunya et al., 2011). The topology of the concatenated
434 28S+18S rRNA tree places the *Cx. portesi* and *Cx. pedroi* species-clades as sister groups (92%
435 bootstrap support), with *Cx. spissipes* as a basal group within the *Melanoconion* clade (100%
436 bootstrap support) (Figure 4, in coral). This corroborates the systematics elucidated by Navarro and
437 Weaver (2004) using the ITS2 marker, and those by Sirivanakarn (1982) and Sallum and Forattini
438 (1996) based on morphology. Curiously, in the COI tree, *Cx. Spissipes* sequences were clustered
439 with unknown species *Cx. Sp1*, forming a clade sister to another containing other *Culex* (*Culex*) and
440 *Culex* (*Oculeomyia*) species, albeit with very low bootstrap support (Figure 5, in coral). Previous

441 phylogenetic studies based on the COI gene have consistently placed *Cx. spissipes* or the Spissipes
442 group basal to other groups within the *Melanoconion* subgenus (Torres-Gutierrez et al., 2016, 2018).
443 However, these studies contain only *Culex (Melanoconion)* species in their assemblage, apart from
444 *Cx. quinquefasciatus* to act as an outgroup. This clustering of *Cx. spissipes* with non-*Melanoconion*
445 species in our COI phylogeny could be an artefact of a much more diversified assemblage rather than
446 a true phylogenetic link.

447 Taking advantage of our multi-country sampling, we examined whether rRNA or COI phylogeny
448 can be used to distinguish conspecifics originating from different geographies. Our assemblage
449 contains five of such species: *An. coustani*, *An. funestus*, *An. gambiae*, *Ae. albopictus*, and *Ma.*
450 *uniformis*. Among the rRNA trees, the concatenated 28S+18S and 28S rRNA trees were able to
451 discriminate between *Ma. uniformis* specimens from Madagascar, Cambodia, and the Central African
452 Republic (in dark green), and between *An. coustani* specimens from Madagascar and the Central
453 African Republic (in purple) (100% bootstrap support). In the COI tree, only *Ma. uniformis* was
454 resolved into geographical clades comprising specimens from Madagascar and specimens from
455 Cambodia (in dark green) (72% bootstrap support). No COI sequence was obtained from one *Ma.*
456 *uniformis* specimen from the Central African Republic. The 28S+18S rRNA sequences ostensibly
457 provided more population-level genetic information than COI sequences alone with better support.
458 The use of rRNA sequences in investigating the biodiversity of mosquitoes should therefore be
459 explored with a more comprehensive taxonomic assemblage.

460 The phylogenetic reconstructions based on rRNA or COI sequences in our study are hardly
461 congruent (Table 2), but two principal differences stand out. First, the COI phylogeny does not
462 recapitulate the early divergence of *Anophelinae* from *Culicinae* (Figure 5). This is at odds with other
463 studies estimating mosquito divergence times based on mitochondrial genes (Logue et al., 2013;
464 Lorenz et al., 2021) or nuclear genes (Reidenbach et al., 2009). The second notable feature in the
465 rRNA trees is the remarkably large interspecies and intersubgeneric evolutionary distances within
466 genus *Anopheles* relative to genera in the *Culicinae* subfamily (Figure 3, Figure 3—figure supplement
467 1, Figure 4, Figure 4—figure supplements 1 and 2; *Anopheles* in purple) but this is not apparent in the
468 COI tree. The hyperdiversity among *Anopheles* taxa may be attributed to the earlier diversification of
469 the *Anophelinae* subfamily in the early Cretaceous period compared to that of the *Culicinae* subfamily,
470 a difference of at least 40 million years (Lorenz et al., 2021). The differences in rRNA and COI tree

471 topologies indicate a limitation in using COI alone to determine evolutionary relationships. Importantly,
472 drawing phylogenetic conclusions from short DNA barcodes such as COI has been cautioned against
473 due to its weak phylogenetic signal (Hajibabaei et al., 2006). The relatively short length of our COI
474 sequences (621–699 bp) combined with the 100-fold higher nuclear substitution rate of mitochondrial
475 genomes relative to nuclear genomes (Arctander, 1995) could result in homoplasy (Danforth et al.,
476 2005), making it difficult to clearly discern ancestral sequences and correctly assign branches into
477 lineages, as evidenced by the poor nodal bootstrap support at genus-level branches. Indeed, in the
478 study by Lorenz et al. (2021), a phylogenetic tree constructed using a concatenation of all 13 protein-
479 coding genes of the mitochondrial genome was able to resolve ancient divergence events. This
480 affirms that while COI sequences can be used to reveal recent speciation events, longer or multi-gene
481 molecular markers are necessary for studies into deeper evolutionary relationships (Danforth et al.,
482 2005).

483 In contrast to Anophelines where 28S rRNA phylogenies illustrated higher interspecies divergence
484 compared to COI phylogeny, two specimens of an unknown *Mansonia* species, “Ma sp.4 GF S103”
485 and “Ma sp.4 GF S104”, provided an example where interspecies relatedness based on their COI
486 sequences is greater than that based on their rRNA sequences in relation to “Ma titillans GF S105”.
487 While all rRNA trees placed “Ma titillans GF S105” as a sister taxon with 100% bootstrap support, the
488 COI tree placed Ma sp.4 basal to all other species except *Ur. Geometrica* (Figure 5; *Mansonia* in dark
489 green, *Uranotaenia* in pink). This may hint at a historical selective sweep in the mitochondrial
490 genome, whether arising from geographical separation, mutations, or linkage disequilibrium with
491 inherited symbionts (Hurst & Jiggins, 2005), resulting in the disparate mitochondrial haplogroups
492 found in French Guyanese Ma sp.4 and *Ma. titillans*. In addition, both haplogroups are distant from
493 those associated with members of subgenus *Mansonoidea*. To note, the COI sequences of “Ma sp.4
494 GF S103” and “Ma sp.4 GF S104” share 87.12 and 87.39% nucleotide similarity, respectively, to that
495 of “Ma titillans GF S105”. Interestingly, the endosymbiont *Wolbachia pipientis* has been detected in
496 *Ma. titillans* sampled from Brazil (De Oliveira et al., 2015), which may contribute to the divergence of
497 “Ma titillans GF S105” COI sequence away from those of Ma sp.4. This highlights other caveats of
498 using a mitochondrial DNA marker in determining evolutionary relationships (Hurst & Jiggins, 2005),
499 which nuclear markers such as 28S and 18S rRNA sequences may be immune to.

500 **Conclusions**

501 Total RNA-seq is a valuable tool for surveillance and virus discovery in sylvatic mosquitoes but it is
 502 impeded by the lack of full-length rRNA reference sequences. Here we presented a rRNA sequence
 503 assembly strategy and a dataset of 234 newly generated mosquito 28S and 18S rRNA sequences.
 504 Our work has expanded the current mosquito rRNA reference library by providing, to our knowledge,
 505 the first full-length rRNA records for 30 species in public databases and paves the way for the
 506 assembly of many more. These novel rRNA sequences can improve mosquito metagenomics based
 507 on RNA-seq by enabling physical and computational removal of rRNA from specimens and
 508 streamlined species identification using rRNA markers.

509 Given that a reference sequence is available, rRNA markers could serve as a better approach for
 510 mosquito taxonomy and phylogeny than COI markers. In analysing the same set of specimens by
 511 their COI and rRNA sequences, we showed that rRNA sequences can discriminate between
 512 members of a species subgroup as well as conspecifics from different geographies. Phylogenetic
 513 inferences from a tree based on 28S rRNA sequences alone or on concatenated 28S+18S rRNA
 514 sequences are more aligned with contemporary mosquito systematics, showing evolutionary
 515 relationships that agree with other phylogenetic studies. While COI-based phylogeny can reveal
 516 recent speciation events, rRNA sequences may be better suited for investigations of deeper
 517 evolutionary relationships as they are less prone to selective sweeps and homoplasy. The
 518 advantages and disadvantages of rRNA and COI sequences as molecular markers are summarised
 519 in Table 3. Further studies are necessary to reveal how rRNA sequences compare against other
 520 nuclear or mitochondrial DNA marker systems (Batovska et al., 2017; Beebe, 2018; Behura, 2006;
 521 Ratnasingham & Hebert, 2007; Reidenbach et al., 2009; Vezenegho et al., 2022).

522 **Table 3.** Comparison of 28S or concatenated 28S+18S rRNA and COI sequences as molecular
 523 markers.

| 28S+18S rRNA | |
|---|---|
| Advantages | Disadvantages |
| <p>In RNA-seq metagenomics studies, molecular taxonomy of specimens based on rRNA sequences can be done from RNA-seq data without additional sample preparation or sequencing.</p> <p>28S rRNA and concatenated 28S+18S rRNA sequences can resolve the identity of specimens where COI sequences were ambiguous, particularly between members of species subgroups.</p> | <p>RNA-seq costs more than Sanger sequencing.</p> <p>Reference rRNA sequences are currently much more limited in breadth compared to other established molecular markers.</p> |

| | |
|--|--|
| <p>28S rRNA and concatenated 28S+18S rRNA sequences can distinguish conspecifics from different geographies for certain species.</p> <p>Phylogenetic inferences based on 28S rRNA and concatenated 28S+18S rRNA sequences show relationships that are more concordant to contemporary mosquito systematics elucidated by other studies and may be a more suitable marker to study deep evolutionary relationships.</p> <p>Being longer and nuclear-encoded, 28S or concatenated 28S+18S rRNA sequences are immune to homoplasy or to selective sweeps that may affect genomes of inherited symbionts such as mitochondria.</p> | |
| COI | |
| Advantages | Disadvantages |
| <p>With a larger reference database, the COI is a versatile marker for molecular taxonomy.</p> <p>Being a shorter DNA marker, the COI gene is cost-and time-effective to amplify, sequence, and characterise.</p> <p>Universal primer sets to amplify the COI marker have been developed and tested for many diverse species.</p> | <p>All species taxa clustered into distinct clades but with weaker bootstrap support at internal nodes relative to those of the 28S+18S rRNA tree.</p> <p>For <i>An. coustani</i>, and members of <i>Culex</i> species subgroups such as <i>Cx. quinquefasciatus</i> and <i>Cx. tritaeniorhynchus</i>, COI is unable to unequivocally confirm species identity as species can differ by just one nucleotide. Other molecular markers are often used in tandem.</p> |

524

525 **MATERIALS AND METHODS**

526 **Sample collection**

527 Mosquito specimens were sampled from 2019 to 2020 by medical entomology teams from the Institut
528 Pasteur de Bangui (Central African Republic, Africa; CF), Institut Pasteur de Madagascar
529 (Madagascar, Africa; MG), Institut Pasteur du Cambodge (Cambodia, Asia; KH), and Institut Pasteur
530 de la Guyane (French Guiana, South America; GF). Adult mosquitoes were sampled using several
531 techniques including CDC light traps, BG sentinels, and human-landing catches. Sampling sites are
532 sylvatic locations including rural settlements in the Central African Republic, Madagascar, and French
533 Guiana and national parks in Cambodia. Mosquitoes were morphologically identified using taxonomic
534 identification keys (Edwards, 1941; Grjebine, 1966; Huang & Ward, 1981; Oo et al., 2006;
535 Rattarithikul et al., 2007, 2010; Rattarithikul, Harbach, et al., 2005; Rattarithikul, Harrison, et
536 al., 2005; Rattarithikul, Harrison, Harbach, et al., 2006; Rattarithikul, Harrison, Panthusiri, et al.,
537 2006; Rueda, 2004) on cold tables before preservation by flash freezing in liquid nitrogen and

538 transportation in dry ice to Institut Pasteur Paris for analysis. A list of the 112 mosquito specimens
539 included in our taxonomic assemblage and their related information are provided in Appendix 1—table
540 1. To note, specimen ID S53, S80, and S81 were removed from our assemblage as their species
541 identity could not be determined by COI or rRNA sequencing.

542 **RNA and DNA isolation**

543 Nucleic acids were isolated from mosquito specimens using TRIzol reagent according to
544 manufacturer's protocol (Invitrogen, Thermo Fisher Scientific, Waltham, Massachusetts, USA). Single
545 mosquitoes were homogenised into 200 μ L of TRIzol reagent and other of the reagents within the
546 protocol were volume-adjusted accordingly. Following phase separation, RNA were isolated from the
547 aqueous phase while DNA were isolated from the remaining interphase and phenol-chloroform phase.
548 From here, RNA is used to prepare cDNA libraries for next generation sequencing while DNA is used
549 in PCR amplification and Sanger sequencing of the mitochondrial *cytochrome c oxidase subunit I*
550 (COI) gene as further described below.

551 **Probe depletion of rRNA**

552 We tested a selective rRNA depletion protocol by Morlan *et al.* (2012) on several mosquito species
553 from the *Aedes*, *Culex*, and *Anopheles* genera. We designed 77 tiled 80 bp DNA probes antisense to
554 the *Ae. aegypti* 28S, 18S, and 5.8S rRNA sequences. A pool of probes at a concentration of 0.04 μ M
555 were prepared. To bind probes to rRNA, 1 μ L of probes and 2 μ L of Hybridisation Buffer (100 mM
556 Tris-HCl and 200 mM NaCl) was added to rRNA samples to a final volume of 20 μ L and subjected to
557 a slow-cool incubation starting at 95 $^{\circ}$ C for 2 minutes, then cooling to 22 $^{\circ}$ C at a rate of 0.1 $^{\circ}$ C per
558 second, ending with an additional 5 minutes at 22 $^{\circ}$ C. The resulting RNA:DNA hybrids were treated
559 with 2.5 μ L Hybridase™ Thermostable RNase H (Epicentre, Illumina, Madison, Wisconsin, USA) and
560 incubated at 37 $^{\circ}$ C for 30 minutes. To remove DNA probes, the mix was treated with 1 μ L DNase I
561 (Invitrogen) and purified with Agencourt RNAClean XP Beads (Beckman Coulter, Brea, California,
562 USA). The resulting RNA is used for total RNA sequencing to check depletion efficiency.

563 **Total RNA sequencing**

564 To obtain rRNA sequences, RNA samples were quantified on a Qubit Fluorometer (Invitrogen) using
565 the Qubit RNA BR Assay kit (Invitrogen) for concentration adjustment. Non-depleted total RNA was
566 used for library preparation for next generation sequencing using the NEBNext Ultra II RNA Library

567 Preparation Kit for Illumina (New England Biolabs, Ipswich, Massachusetts, USA) and the NEBNext
568 Multiplex Oligos for Illumina (Dual Index Primers Set 1) (New England Biolabs). Sequencing was
569 performed on a NextSeq500 sequencing system (Illumina, San Diego, California, USA). Quality
570 control of fastq data and trimming of adapters were performed with FastQC and cutadapt,
571 respectively.

572 **28S and 18S rRNA assembly**

573 To obtain 28S and 18S rRNA contigs, we had to first clean our fastq library by separating the reads
574 representing mosquito rRNA from all other reads. To achieve this, we used the SILVA RNA sequence
575 database to create 2 libraries: one containing all rRNA sequences recorded under the "Insecta" node
576 of the taxonomic tree, the other containing the rRNA sequences of many others nodes distributed
577 throughout the taxonomic tree, hence named "Non-Insecta" (Quast et al., 2013). Each read was
578 aligned using the nucleotide Basic Local Alignment Search Tool (BLASTn,
579 <https://blast.ncbi.nlm.nih.gov/>) of the National Center for Biotechnology Information (NCBI) against
580 each of the two libraries and the scores of the best high-scoring segment pairs from the two BLASTns
581 are subsequently used to calculate a ratio of Insecta over Non-Insecta scores (Altschul et al., 1990).
582 Only reads with a ratio greater than 0.8 were used in the assembly. The two libraries being non-
583 exhaustive, we chose this threshold of 0.8 to eliminate only reads that were clearly of a non-insect
584 origin. Selected reads were assembled with the SPAdes genome assembler using the "-rna" option,
585 allowing more heterogeneous coverage of contigs and kmer lengths of 31 , 51 and 71 bases
586 (Bankevich et al., 2012). This method successfully assembled rRNA sequences for all specimens,
587 including a parasitic *Horreolanus* water mite (122 sequences for 28S and 114 sequences for 18S).

588 Initially, our filtration technique had two weaknesses. First, there is a relatively small number of
589 complete rRNA sequences in the Insecta library from SILVA. To compensate for this, we carried out
590 several filtration cycles, each time adding in the complete sequences produced in previous cycles to
591 the Insecta library. Second, when our mosquito specimens were parasitized by other insects, it was
592 not possible to bioinformatically filter out rRNA reads belonging to the parasite. For these rare cases,
593 we used the "--trusted-contigs" option of the SPAdes assembler (Bankevich et al., 2012), giving it
594 access to the 28S and 18S rRNA sequences of the mosquito closest in terms of taxonomic distance.
595 By doing this, the assembler was able to reconstruct the rRNA of the mosquito as well as the rRNA of
596 the parasitizing insect. All assembled rRNA sequences from this study have been deposited in

597 GenBank with accession numbers OM350214–OM350327 for 18S rRNA sequences and OM542339–
598 OM542460 for 28S rRNA sequences.

599 **COI amplicon sequencing**

600 The mitochondrial COI gene was amplified from DNA samples using the universal “Folmer” primer set
601 LCO1490 (5'- GGTCACAAATCATAAAGATATTGG -3') and HCO2198 (5'-
602 TAAACTTCAGGGTGACCAAAAAATCA-3'), as per standard COI marker sequencing practices,
603 producing a 658 bp product (Folmer et al., 1994). PCRs were performed using Phusion High-Fidelity
604 DNA Polymerase (Thermo Fisher Scientific). Every 50 µL reaction contained 10 µL of 5X High Fidelity
605 buffer, 1 µL of 10 mM dNTPs, 2.5 µL each of 10 mM forward (LCO1490) and reverse (HCO2198)
606 primer, 28.5 µL of water, 5 µL of DNA sample, and 0.5 µL of 2 U/µL Phusion DNA polymerase. A 3-
607 step cycling incubation protocol was used: 98 °C for 30 seconds; 35 cycles of 98 °C for 10 seconds,
608 60 °C for 30 seconds, and 72 °C for 15 seconds; 72 °C for 5 minutes ending with a 4 °C hold. PCR
609 products were size-verified using gel electrophoresis and then gel-purified using the QIAquick Gel
610 Extraction Kit (Qiagen, Hilden, Germany). Sanger sequencing of the COI amplicons were performed
611 by Eurofins Genomics, Ebersberg, Germany.

612 **COI sequence analysis**

613 Forward and reverse COI DNA sequences were end-trimmed to remove bases of poor quality (Q
614 score < 30). At the 5' ends, sequences were trimmed at the same positions such that all forward
615 sequences start with 5'- TTTTGG and all reverse sequences start with 5'- GGNTCT. Forward and
616 reverse sequences were aligned using BLAST to produce a 621 bp consensus sequence. In cases
617 where good quality sequences extends beyond 621 bp, forward and reverse sequences were
618 assembled using Pearl (<https://www.gear-genomics.com/pearl/>) and manually checked for errors
619 against trace files (Rausch et al., 2019, 2020). We successfully assembled a total of 106 COI
620 sequences. All assembled COI sequences from this study have been deposited in GenBank with
621 accession numbers OM630610–OM630715.

622 **COI validation of morphology-based species identification**

623 We analysed assembled COI sequences with BLASTn against the nucleotide collection (nr/nt)
624 database to confirm morphology-based species identification. BLAST analyses revealed 32 cases
625 where top hits indicated a different species identity, taking <95% nucleotide sequence similarity as the

626 threshold to delineate distinct species (Appendix 2—table 1). In these cases, the COI sequence of the
627 specimen was then BLAST-aligned against a GenBank record representing the morphological
628 species to verify that the revised identity is a closer match by a significant margin, i.e., more than 2%
629 nucleotide sequence similarity. All species names reported hereafter reflect identities determined by
630 COI sequence except for cases where COI-based identities were ambiguous, in which case
631 morphology-based identities were retained. In cases where matches were found within a single genus
632 but of multiple species, specimens were indicated as an unknown member of their genus (e.g., *Culex*
633 sp.). Information of the highest-scoring references for all specimens, including details of ambiguous
634 BLASTn results, are recorded in Appendix 2—table 1.

635 Within our COI sequences, we found six unidentified *Culex* species (including two that matched to
636 GenBank entries identified only to the genus level), four unidentified *Mansonia* species, and one
637 unidentified *Mimomyia* species. For *An. baezai*, no existing GenBank records were found at the time
638 this analysis was performed.

639 **Phylogenetic analysis**

640 Multiple sequence alignment (MSA) were performed on assembled COI and rRNA sequences using
641 the MUSCLE software (Edgar, 2004; Madeira et al., 2019). As shown in Figure 3—figure supplement
642 2, the 28S rRNA sequences contain many blocks of highly conserved nucleotides, which makes the
643 result of multiple alignment particularly obvious. We therefore did not test other alignment programs.
644 The multiple alignment of the COI amplicons is even more evident since no gaps are necessary for
645 this alignment.

646 Phylogenetic tree reconstructions were performed with the MEGA X software using the maximum-
647 likelihood method (S. Kumar et al., 2018). Default parameters were used with bootstrapping with 500
648 replications to quantify confidence level in branches. For rRNA trees, sequences belonging to an
649 unknown species of parasitic water mite (genus *Horreolanus*) found in our specimens served as an
650 outgroup taxon. In addition, we created and analysed a separate dataset combining our 28S rRNA
651 sequences and full-length 28S rRNA sequences from GenBank totalling 169 sequences from 58
652 species (12 subgenera). To serve as outgroups for the COI tree, we included sequences obtained
653 from GenBank of three water mite species, *Horreolanus orphanus* (KM101004), *Sperchon fuxiensis*
654 (MH916807), and *Arrenurus* sp. (MN362807).

655 **DECLARATIONS**

656 **Availability of data and materials**

657 The multiple sequence alignments that gave rise to the presented phylogenetic reconstructions are
658 included as source data files linked to the relevant figures. All sequences generated in this study have
659 been deposited in GenBank under the accession numbers OM350214–OM350327 for 18S rRNA
660 sequences, OM542339–OM542460 for 28S rRNA sequences, and OM630610–OM630715 for COI
661 sequences (Appendix 1—table 1).

662 **Competing interests**

663 The authors declare that they have no financial or non-financial competing interests.

664 **Impact statement**

665 A score-based read filtration strategy enables assembly of full-length ribosomal RNA sequences for
666 lesser-studied mosquito species, opening the doors to the use of these sequences as a novel
667 molecular marker for taxonomic and phylogenetic studies.

668 **Funding**

669 This work was supported by the Defence Advanced Research Projects Agency PREEMPT program
670 managed by Dr. Rohit Chitale and Dr. Kerri Dugan [Cooperative Agreement HR001118S0017] (the
671 content of the information does not necessarily reflect the position or the policy of the U.S.
672 government, and no official endorsement should be inferred).

673 **Acknowledgements**

674 We thank members of the Saleh lab for valuable discussions and to Dr Louis Lambrechts for critical
675 reading of the manuscript. We especially thank all medical entomology staff of IP Bangui, IP
676 Cambodge (Sony Yean, Kimly Heng, Kalyan Chhuoy, Sreynik Nhek, Moeun Chhum, Kimhuor Sour
677 and Pierre-Olivier Maquart), IP Madagascar, and IP Guyane for assistance in field missions,
678 laboratory work, and logistics, and Inès Partouche from IP Paris for laboratory assistance. We are
679 also grateful to Dr Catherine Dauga for advice on phylogenetic analyses, and to Amandine Guidez for
680 providing a French Guiana-specific COI reference library. Finally, we thank our reviewers, including
681 Leslie Vosshall and Katherine Young, for their constructive reviews of this manuscript.

682 **FIGURE SUPPLEMENTS**

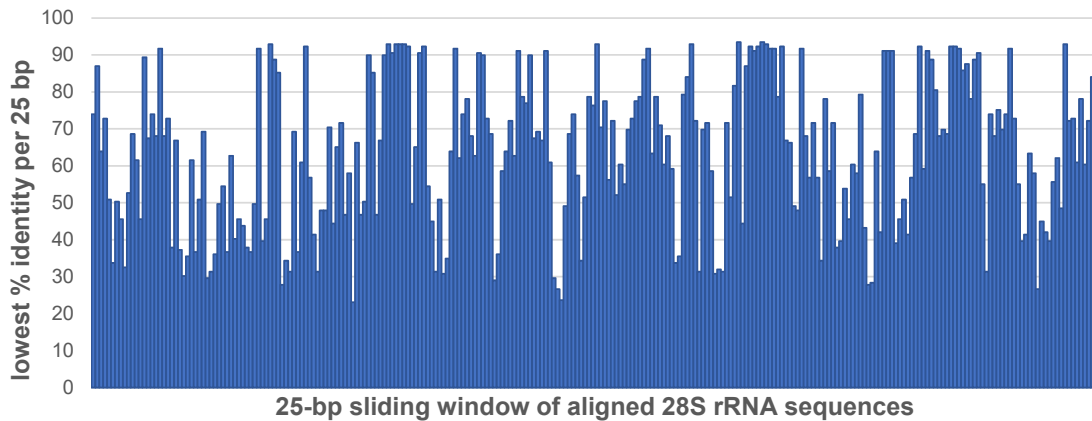
683 **Figure 3—figure supplement 1**

684

685 **Interspecific and intersubgeneric distances within the genus *Anopheles* indicate a greater**
686 **degree of divergence than those within any other genera of family *Culicidae*.** The phylogenetic
687 tree presented in Figure 3 based on 28S sequences from this study and from GenBank (annotated
688 with filled circles) is depicted here in radial format to illustrate how the branch lengths separating
689 Anopheline taxa are longer relative to other members of family *Culicidae*. An unknown Horreolanus

690 species found among our samples serves as an outgroup. For sequences from this study, each
691 specimen label contains information on taxonomy, origin (in 2-letter country codes), and specimen ID
692 number. Specimen genera are indicated by colour: *Culex* in coral, *Anopheles* in purple, *Aedes* in dark
693 blue, *Mansonia* in dark green, *Limatus* in light green, *Coquillettidia* in light blue, *Psorophora* in yellow,
694 *Mimomyia* in teal, *Uranotaenia* in pink and *Eretmapodites* in brown. Scale bar at 0.05 is shown.

695 **Figure 3—figure supplement 2**



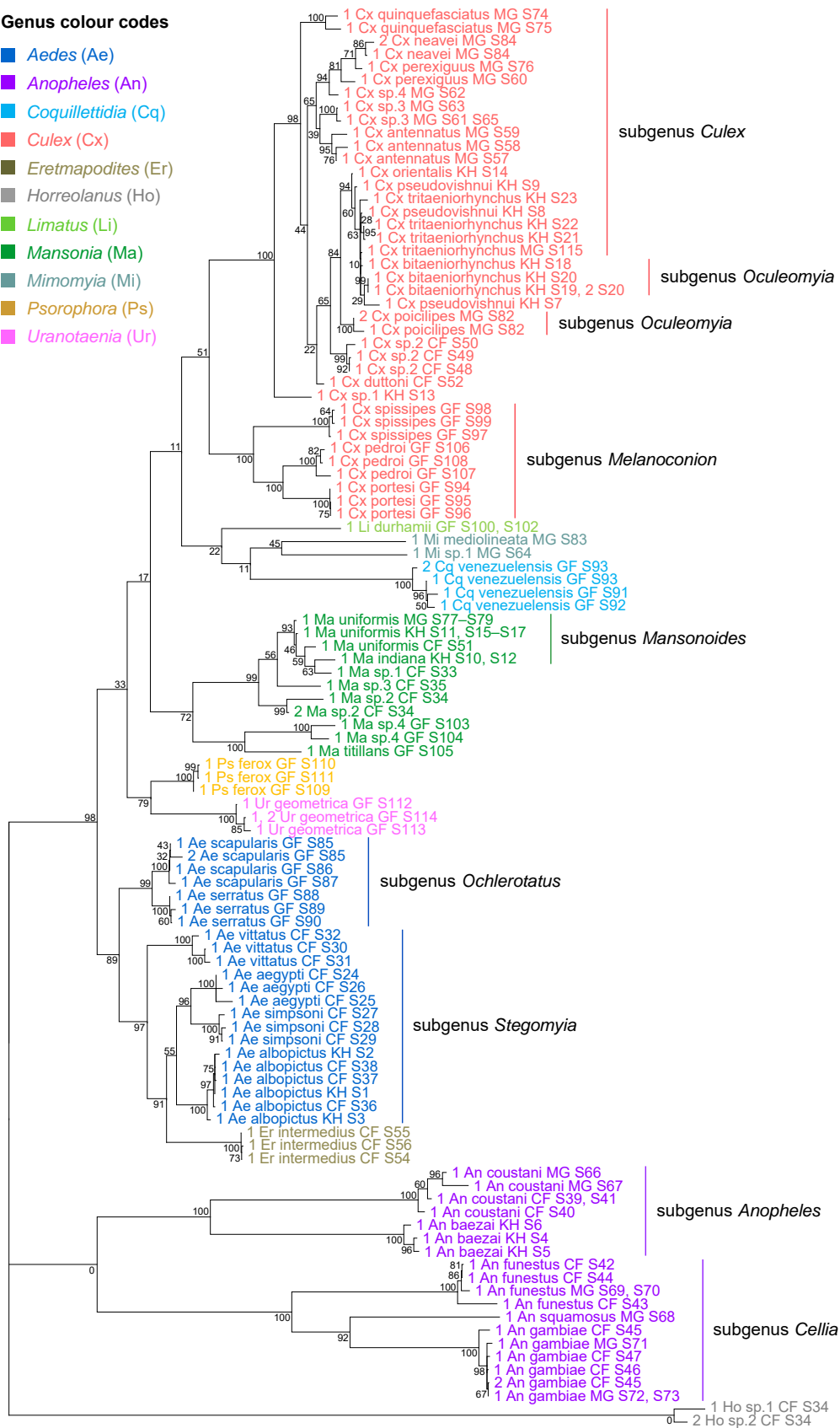
696

697 **Sequence conservation among 169 28S rRNA sequences obtained from this study and from**
698 **GenBank combined.** Multiple sequence alignment was performed on 28S rRNA sequences, 3900 bp
699 in length. Each bar represents a 25-bp sliding window of the 28S rRNA sequence alignment where
700 the y-axis values are the lowest percentage nucleotide identity found.

701 **Figure 4—figure supplement 1**

Genus colour codes

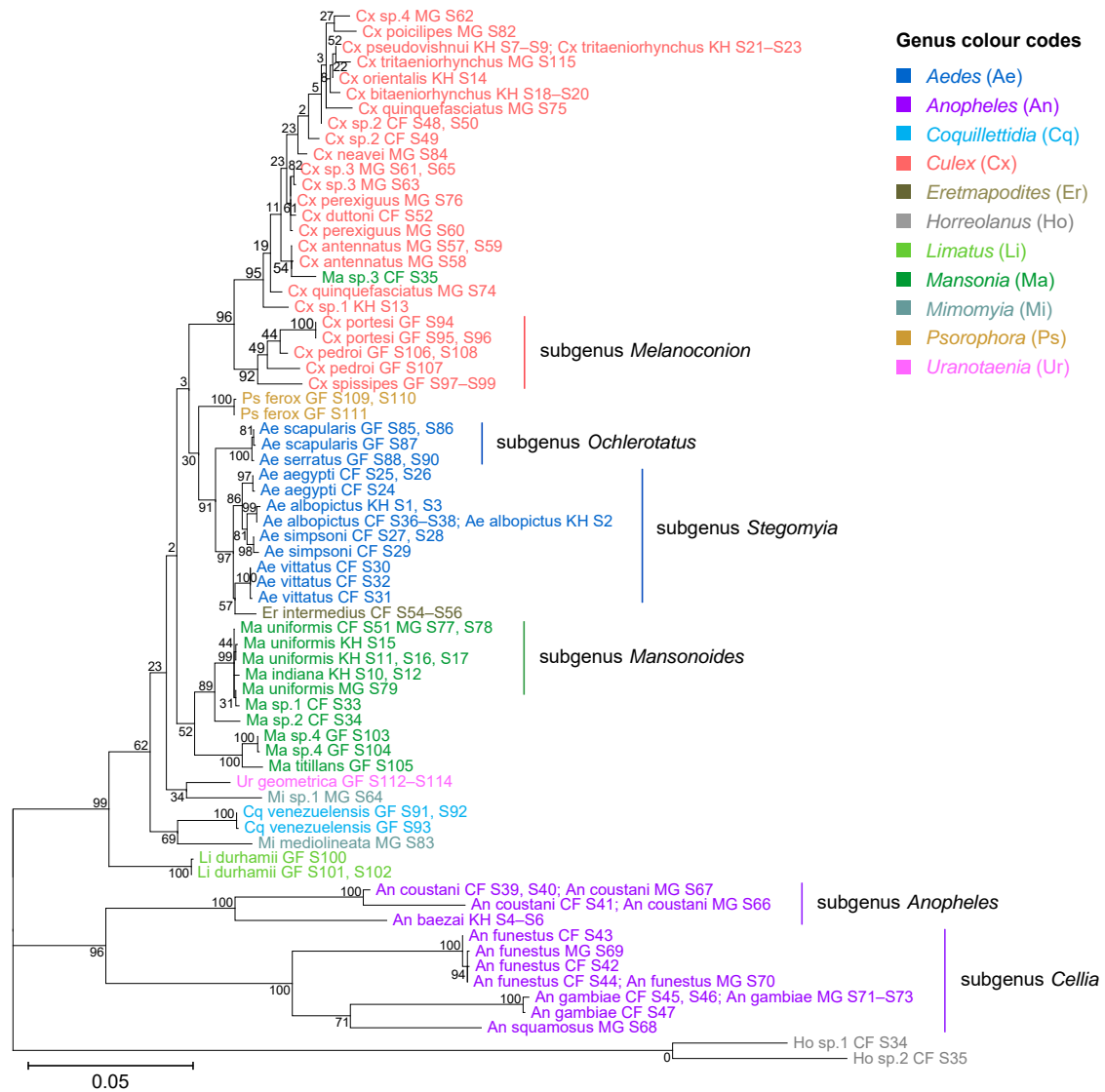
- *Aedes* (Ae)
- *Anopheles* (An)
- *Coquillettidia* (Cq)
- *Culex* (Cx)
- *Eretmapodites* (Er)
- *Horreolanus* (Ho)
- *Limatus* (Li)
- *Mansonia* (Ma)
- *Mimomyia* (Mi)
- *Psorophora* (Ps)
- *Uranotaenia* (Ur)



0.05

703 **Phylogenetic tree based on 28S rRNA sequences generated from this study (3900 bp).** This tree
704 was inferred using maximum-likelihood method and constructed to scale in MEGA X (S. Kumar et al.,
705 2018) using an unknown *Horreolanus* species found among our samples as an outgroup. Values at
706 each node indicate bootstrap support (%) from 500 replications. For sequences from this study, each
707 specimen label contains information on taxonomy, origin (in 2-letter country codes), and specimen ID
708 number. Some specimens produced up to two consensus 28S rRNA sequences; this is indicated by
709 the numbers 1 or 2 in the beginning of the specimen label. Specimen genera are indicated by colour:
710 *Culex* in coral, *Anopheles* in purple, *Aedes* in dark blue, *Mansonia* in dark green, *Limatus* in light
711 green, *Coquillettidia* in light blue, *Psorophora* in yellow, *Mimomyia* in teal, *Uranotaenia* in pink and
712 *Eretmapodites* in brown. Scale bar at 0.05 is shown.

713 **Figure 4—figure supplement 2**



714

715 **Phylogenetic tree based on 18S rRNA sequences (1900 bp).** This tree was inferred using
 716 maximum-likelihood method and constructed to scale in MEGA X (S. Kumar et al., 2018) using an
 717 unknown *Horreolanus* species found among our samples as an outgroup. Values at each node
 718 indicate bootstrap support (%) from 500 replications. For sequences from this study, each specimen
 719 label contains information on taxonomy, origin (in 2-letter country codes), and specimen ID number.
 720 One 18S rRNA sequence was obtain for each specimen. Specimen genera are indicated by colour:
 721 *Culex* in coral, *Anopheles* in purple, *Aedes* in dark blue, *Mansonia* in dark green, *Limatus* in light
 722 green, *Coquillettidia* in light blue, *Psorophora* in yellow, *Mimomyia* in teal, *Uranotaenia* in pink and
 723 *Eretmapodites* in brown. Scale bar at 0.05 is shown.

724 APPENDICES

725 **Appendix 1—table 1.** Taxonomic and sampling information on mosquito specimens and associated accession numbers of their COI, 18S rRNA, and 28S
726 rRNA sequences (XLSX).

| Sequence ID | Taxonomy [Genus (subgenus) species] | Origin | Collection site | Collection period | Blood engorged (Y/N) | Sample ID | COI accession number | 18S rRNA accession number | 28S rRNA accession number |
|------------------------|--|----------|-----------------|-------------------|-------------------------|-----------|----------------------|---------------------------|---------------------------|
| Ae_albopictus_KH_S1 | Aedes (Stegomyia) albopictus | Cambodia | Rattanakiri | Dec 2019 | N | 1 | OM630613 | OM350214 | OM542460 |
| Ae_albopictus_KH_S2 | Aedes (Stegomyia) albopictus | Cambodia | Rattanakiri | Dec 2019 | N | 2 | OM630614 | OM350220 | OM542373 |
| Ae_albopictus_KH_S3 | Aedes (Stegomyia) albopictus | Cambodia | Rattanakiri | Dec 2019 | N | 3 | OM630615 | OM350316 | OM542374 |
| An_baezai_KH_S4 | Anopheles (Anopheles) baezai | Cambodia | Koh Kong | Mar 2019 | N | 4 | OM630631 | OM350327 | OM542357 |
| An_baezai_KH_S5 | Anopheles (Anopheles) baezai | Cambodia | Koh Kong | Mar 2019 | N | 5 | OM630632 | OM350233 | OM542440 |
| An_baezai_KH_S6 | Anopheles (Anopheles) baezai | Cambodia | Koh Kong | Mar 2019 | N | 6 | OM630633 | OM350234 | OM542358 |
| Cx_pseudovishnui_KH_S7 | Culex (Culex) pseudovishnui | Cambodia | Rattanakiri | Dec 2019 | N | 7 | OM630689 | OM350285 | OM542413 |
| Cx_pseudovishnui_KH_S8 | Culex (Culex) pseudovishnui | Cambodia | Rattanakiri | Dec 2019 | N | 8 | OM630690 | OM350286 | OM542414 |
| Cx_pseudovishnui_KH_S9 | Culex (Culex) pseudovishnui | Cambodia | Rattanakiri | Dec 2019 | N | 9 | OM630691 | OM350287 | OM542415 |
| Ma_indiana_KH_S10 | Mansonia (Mansonioides) indiana | Cambodia | Battambang | Nov 2019 | N | 10 | OM630698 | OM350295 | OM542422 |
| Ma_uniformis_KH_S11 | Mansonia (Mansonioides) uniformis | Cambodia | Battambang | Nov 2019 | N | 11 | OM630699 | OM350296 | OM542423 |
| Ma_indiana_KH_S12 | Mansonia | Cambodia | Battambang | Nov 2019 | N | 12 | OM630700 | OM350297 | OM542424 |

| | | | | | | | | | | |
|-----------------------------|---|--------------------------------|------------|----------|---|----|----------|----------|-----------------------|--|
| | (Mansonioides) indiana | | | | | | | | | |
| Cx_sp.1_KH_S13 | Culex sp.1 Culex (Culex) | Cambodia | Prek Toal | Feb 2019 | N | 13 | OM630672 | OM350267 | OM542395 | |
| Cx_orientalis_KH_S14 | orientalis Mansonia (Mansonioides) | Cambodia | Prek Toal | Feb 2019 | N | 14 | OM630673 | OM350268 | OM542396 | |
| Ma_uniformis_KH_S15 | uniformis Mansonia (Mansonioides) | Cambodia | Battambang | Nov 2019 | N | 15 | OM630705 | OM350303 | OM542430 | |
| Ma_uniformis_KH_S16 | uniformis Mansonia (Mansonioides) | Cambodia | Battambang | Nov 2019 | N | 16 | OM630706 | OM350305 | OM542432 | |
| Ma_uniformis_KH_S17 | uniformis Culex (Oculeomyia) | Cambodia | Battambang | Nov 2019 | N | 17 | OM630707 | OM350304 | OM542431 | |
| Cx_bitaeniorhynchus_KH_S18 | bitaeniorhynchus Culex (Oculeomyia) | Cambodia | Battambang | Nov 2019 | N | 18 | OM630656 | OM350255 | OM542381 | |
| Cx_bitaeniorhynchus_KH_S19 | bitaeniorhynchus Culex (Oculeomyia) | Cambodia | Battambang | Nov 2019 | N | 19 | OM630657 | OM350256 | OM542382 | |
| Cx_bitaeniorhynchus_KH_S20 | bitaeniorhynchus Culex (Culex) | Cambodia | Battambang | Nov 2019 | N | 20 | OM630658 | OM350257 | OM542383, OM542384 | |
| Cx_tritaeniorhynchus_KH_S21 | tritaeniorhynchus Culex (Culex) | Cambodia | Battambang | Nov 2019 | N | 21 | OM630680 | OM350277 | OM542404 | |
| Cx_tritaeniorhynchus_KH_S22 | tritaeniorhynchus Culex (Culex) | Cambodia | Battambang | Nov 2019 | N | 22 | OM630681 | OM350278 | OM542405 | |
| Cx_tritaeniorhynchus_KH_S23 | tritaeniorhynchus Aedes (Stegomyia) | Cambodia | Battambang | Nov 2019 | N | 23 | OM630682 | OM350279 | OM542406 | |
| Ae_aegypti_CF_S24 | aegypti Aedes (Stegomyia) | Central African Republic | Pissa | Jun 2019 | N | 24 | OM630610 | OM350314 | OM542339 | |
| Ae_aegypti_CF_S25 | aegypti Aedes (Stegomyia) | Central African Republic | Pissa | Jun 2019 | N | 25 | OM630611 | OM350215 | OM542340 | |
| Ae_aegypti_CF_S26 | aegypti Aedes (Stegomyia) | Central African Republic | Pissa | Jun 2019 | N | 26 | OM630612 | OM350216 | OM542341 | |
| Ae_simpsoni_CF_S27 | simpsoni Aedes (Stegomyia) | Central African Republic | Pissa | Jun 2019 | N | 27 | OM630619 | OM350221 | OM542345 | |
| Ae_simpsoni_CF_S28 | (Stegomyia) | Central African | Pissa | Jun 2019 | N | 28 | OM630620 | OM350222 | OM542346 | |

| | | | | | | | | | |
|----------------------|--|--|---------|----------|---|----|----------|----------|-----------------------|
| Ae_simpsoni_CF_S29 | simpsoni Aedes (Stegomyia) simpsoni | Republic Central African Republic | Pissa | Jun 2019 | N | 29 | OM630621 | OM350223 | OM542347 |
| Ae_vittatus_CF_S30 | Aedes (Fredwardsius) vittatus | African Republic Central | Gbozo | Aug 2019 | Y | 30 | OM630628 | OM350230 | OM542439 |
| Ae_vittatus_CF_S31 | Aedes (Fredwardsius) vittatus | African Republic Central | Gbozo | Aug 2019 | N | 31 | OM630629 | OM350231 | OM542355 |
| Ae_vittatus_CF_S32 | Aedes (Fredwardsius) vittatus | African Republic Central | Gbozo | Aug 2019 | N | 32 | OM630630 | OM350232 | OM542356 |
| Ma_sp.1_CF_S33 | Mansonia sp.1 | African Republic Central | Bayanga | Nov 2019 | Y | 33 | N/A | OM350294 | OM542449 |
| Ma_sp.2_CF_S34 | Mansonia sp.2 | African Republic Central | Bayanga | Nov 2019 | Y | 34 | N/A | OM350322 | OM542450, OM542456 |
| Ho_sp.1_CF_S34 | Horreolanus sp.1 | African Republic Central | Bayanga | Nov 2019 | - | 34 | N/A | OM350325 | OM542457 |
| Ho_sp.2_CF_S34 | Horreolanus sp.2 | African Republic Central | Bayanga | Nov 2019 | - | 34 | N/A | OM350326 | OM542458 |
| Ma_sp.3_CF_S35 | Mansonia sp.3 | African Republic Central | Bayanga | Nov 2019 | Y | 35 | N/A | OM350323 | OM542451 |
| Ae_albopictus_CF_S36 | Aedes (Stegomyia) albopictus | African Republic Central | Pissa | Jun 2019 | N | 36 | OM630616 | OM350217 | OM542342 |
| Ae_albopictus_CF_S37 | Aedes (Stegomyia) albopictus | African Republic Central | Pissa | Jun 2019 | N | 37 | OM630617 | OM350218 | OM542343 |
| Ae_albopictus_CF_S38 | Aedes (Stegomyia) albopictus | African Republic Central | Pissa | Jun 2019 | N | 38 | OM630618 | OM350219 | OM542344 |
| An_coustani_CF_S39 | Anopheles (Anopheles) coustani | African Republic Central | Pissa | Jan 2020 | N | 39 | OM630634 | OM350235 | OM542359 |
| An_coustani_CF_S40 | Anopheles (Anopheles) coustani | African Republic | Pissa | Jan 2020 | N | 40 | OM630635 | OM350236 | OM542360 |

| | | | | | | | | | |
|-----------------------|---|---|---------|----------|---|----|----------|----------|-----------------------|
| An_coustani_CF_S41 | Anopheles (Anopheles) coustani | Central African Republic Central | Pissa | Jan 2020 | N | 41 | OM630636 | OM350237 | OM542361 |
| An_funestus_CF_S42 | Anopheles (Cellia) funestus | African Republic Central | Pissa | Jun 2019 | Y | 42 | OM630640 | OM350241 | OM542365 |
| An_funestus_CF_S43 | Anopheles (Cellia) funestus | African Republic Central | Pissa | Jun 2019 | Y | 43 | OM630641 | OM350242 | OM542366 |
| An_funestus_CF_S44 | Anopheles (Cellia) funestus | African Republic Central | Pissa | Jun 2019 | Y | 44 | OM630642 | OM350243 | OM542367 |
| An_gambiae_CF_S45 | Anopheles (Cellia) gambiae | African Republic Central | Pissa | Jun 2019 | Y | 45 | OM630645 | OM350245 | OM542369, OM542370 |
| An_gambiae_CF_S46 | Anopheles (Cellia) gambiae | African Republic Central | Pissa | Jun 2019 | Y | 46 | OM630646 | OM350246 | OM542371 |
| An_gambiae_CF_S47 | Anopheles (Cellia) gambiae | African Republic Central | Pissa | Jun 2019 | Y | 47 | N/A | OM350247 | OM542372 |
| Cx_sp.2_CF_S48 | Culex sp.2 | African Republic Central | Bayanga | Nov 2019 | Y | 48 | OM630669 | OM350269 | OM542446 |
| Cx_sp.2_CF_S49 | Culex sp.2 | African Republic Central | Bayanga | Nov 2019 | Y | 49 | OM630670 | OM350315 | OM542397 |
| Cx_sp.2_CF_S50 | Culex sp.2 | African Republic Central | Bayanga | Nov 2019 | Y | 50 | OM630671 | OM350270 | OM542398 |
| Ma_uniformis_CF_S51 | Mansonia (Mansonioides) uniformis | African Republic Central | Bouar | May 2019 | Y | 51 | N/A | OM350301 | OM542428 |
| Cx_duttoni_CF_S52 | Culex (Culex) duttoni | African Republic Central | Mbaiki | Jan 2019 | Y | 52 | OM630704 | OM350302 | OM542429 |
| Er_intermedius_CF_S54 | Eretmapodites intermedius | African Republic Central | Pissa | Jun 2019 | N | 54 | OM630692 | OM350288 | OM542416 |
| Er_intermedius_CF_S55 | Eretmapodites intermedius | African Republic | Pissa | Jun 2019 | N | 55 | OM630693 | OM350289 | OM542417 |
| Er_intermedius_CF_S56 | Eretmapodites | Central | Pissa | Jun 2019 | N | 56 | OM630694 | OM350290 | OM542418 |

| | | | | | | | | | | |
|----------------------------|--------------------------------|------------------|----------------|----------|---|----|----------|----------|----------|--|
| | intermedius | African Republic | | | | | | | | |
| Cx_antennatus_MG_S57 | Culex (Culex) antennatus | Madagascar | Ambato Boeny | Feb 2019 | N | 57 | OM630653 | OM350253 | OM542379 | |
| Cx_antennatus_MG_S58 | Culex (Culex) antennatus | Madagascar | Ambato Boeny | Feb 2019 | N | 58 | OM630654 | OM350319 | OM542444 | |
| Cx_antennatus_MG_S59 | Culex (Culex) antennatus | Madagascar | Ambato Boeny | Feb 2019 | N | 59 | OM630655 | OM350254 | OM542380 | |
| Cx_perexiguus_MG_S60 | Culex (Culex) perexiguus | Madagascar | Amparafaravola | Feb 2019 | N | 60 | OM630660 | OM350258 | OM542386 | |
| Cx_sp.3_MG_S61 | Culex sp.3 | Madagascar | Ambato Boeny | Aug 2019 | N | 61 | OM630661 | OM350259 | OM542387 | |
| Cx_sp.4_MG_S62 | Culex sp.4 | Madagascar | Ambato Boeny | Aug 2019 | N | 62 | OM630662 | OM350260 | OM542388 | |
| Cx_sp.3_MG_S63 | Culex sp.3 | Madagascar | Ambato Boeny | Feb 2019 | N | 63 | OM630686 | OM350282 | OM542410 | |
| Mi_sp.1_MG_S64 | Mimomyia sp.1 | Madagascar | Ambato Boeny | Feb 2019 | N | 64 | OM630687 | OM350283 | OM542411 | |
| Cx_sp.3_MG_S65 | Culex sp.3 | Madagascar | Ambato Boeny | Feb 2019 | N | 65 | OM630688 | OM350284 | OM542412 | |
| An_coustani_MG_S66 | Anopheles (Anopheles) coustani | Madagascar | Ambato Boeny | Feb 2019 | N | 66 | OM630637 | OM350238 | OM542362 | |
| An_coustani_MG_S67 | Anopheles (Anopheles) coustani | Madagascar | Ambato Boeny | Feb 2019 | N | 67 | OM630638 | OM350239 | OM542363 | |
| An_squamosus_MG_S68 | Anopheles (Cellia) squamosus | Madagascar | Ambato Boeny | Feb 2019 | N | 68 | OM630639 | OM350240 | OM542364 | |
| An_funestus_MG_S69 | Anopheles (Cellia) funestus | Madagascar | Ambato Boeny | Feb 2019 | N | 69 | OM630643 | OM350244 | OM542368 | |
| An_funestus_MG_S70 | Anopheles (Cellia) funestus | Madagascar | Ambato Boeny | Feb 2020 | N | 70 | OM630644 | OM350317 | OM542441 | |
| An_gambiae_MG_S71 | Anopheles (Cellia) gambiae | Madagascar | Ambato Boeny | Feb 2019 | N | 71 | OM630647 | OM350249 | OM542442 | |
| An_gambiae_MG_S72 | Anopheles (Cellia) gambiae | Madagascar | Ambato Boeny | Feb 2019 | N | 72 | OM630648 | OM350248 | OM542443 | |
| An_gambiae_MG_S73 | Anopheles (Cellia) gambiae | Madagascar | Ambato Boeny | Feb 2019 | N | 73 | OM630649 | OM350318 | OM542459 | |
| Cx_quinquefasciatus_MG_S74 | Culex (Culex) quinquefasciatus | Madagascar | Amparafaravola | Feb 2019 | N | 74 | OM630674 | OM350271 | OM542399 | |
| Cx_quinquefasciatus_MG_S75 | Culex (Culex) quinquefasciatus | Madagascar | Amparafaravola | Feb 2019 | N | 75 | OM630675 | OM350272 | OM542447 | |
| Cx_perexiguus_MG_S76 | Culex (Culex) perexiguus | Madagascar | Mampikony | Aug 2019 | N | 76 | OM630676 | OM350273 | OM542400 | |

| | | | | | | | | | |
|-------------------------|---|------------------|-----------------------|----------|---|----|----------|----------|-----------------------|
| Ma_uniformis_MG_S77 | Mansonia (Mansonioides) uniformis | Madagascar | Ambato Boeny | Feb 2019 | N | 77 | OM630708 | OM350306 | OM542433 |
| Ma_uniformis_MG_S78 | Mansonia (Mansonioides) uniformis | Madagascar | Ambato Boeny | Feb 2019 | N | 78 | OM630709 | OM350307 | OM542434 |
| Ma_uniformis_MG_S79 | Mansonia (Mansonioides) uniformis | Madagascar | Ambato Boeny | Feb 2019 | N | 79 | OM630710 | OM350308 | OM542435 |
| Cx_poecilipes_MG_S82 | Culex poecilipes | Madagascar | Mampikony | Feb 2019 | N | 82 | OM630659 | OM350320 | OM542385, OM542445 |
| Mi_mediolineata_MG_S83 | Mimomyia mediolineata | Madagascar | Ambato Boeny | Feb 2019 | N | 83 | OM630683 | OM350280 | OM542407 |
| Cx_neavei_MG_S84 | Culex (Culex) neavei | Madagascar | Ambato Boeny | Feb 2019 | N | 84 | OM630684 | OM350281 | OM542408, OM542409 |
| Ae_scapularis_GF_S85 | Aedes (Ochlerotatus) scapularis | French Guiana | Hameau Prefontaine | Jul 2019 | N | 85 | OM630624 | OM350224 | OM542348, OM542349 |
| Ae_scapularis_GF_S86 | Aedes (Ochlerotatus) scapularis | French Guiana | Hameau Prefontaine | Jul 2019 | N | 86 | OM630622 | OM350225 | OM542350 |
| Ae_scapularis_GF_S87 | Aedes (Ochlerotatus) scapularis | French Guiana | Hameau Prefontaine | Jul 2019 | N | 87 | OM630623 | OM350226 | OM542351 |
| Ae_serratus_GF_S88 | Aedes (Ochlerotatus) serratus | French Guiana | Hameau Prefontaine | Nov 2020 | N | 88 | OM630625 | OM350227 | OM542352 |
| Ae_serratus_GF_S89 | Aedes (Ochlerotatus) serratus | French Guiana | Hameau Prefontaine | Nov 2020 | N | 89 | OM630626 | OM350228 | OM542353 |
| Ae_serratus_GF_S90 | Aedes (Ochlerotatus) serratus | French Guiana | Hameau Prefontaine | Nov 2020 | N | 90 | OM630627 | OM350229 | OM542354 |
| Cq_venezuelensis_GF_S91 | Coquillettidia venezuelensis | French Guiana | Hameau Prefontaine | Jul 2019 | N | 91 | OM630650 | OM350250 | OM542375 |
| Cq_venezuelensis_GF_S92 | Coquillettidia venezuelensis | French Guiana | Hameau Prefontaine | Jul 2019 | N | 92 | OM630651 | OM350251 | OM542376 |
| Cq_venezuelensis_GF_S93 | Coquillettidia venezuelensis | French Guiana | Hameau Prefontaine | Jul 2019 | N | 93 | OM630652 | OM350252 | OM542377, OM542378 |
| Cx_portesi_GF_S94 | Culex sp. BTLHVDV-2014 | French Guiana | Hameau Prefontaine | Jul 2019 | N | 94 | OM630666 | OM350264 | OM542392 |
| Cx_portesi_GF_S95 | Culex sp. BTLHVDV-2014 | French Guiana | Hameau Prefontaine | Jul 2019 | N | 95 | OM630667 | OM350265 | OM542393 |

| | | | | | | | | | |
|-----------------------|--|------------------|-----------------------|----------|---|-----|----------|----------|----------|
| Cx_portesi_GF_S96 | Culex sp. BTLHVDV-2014 | French Guiana | Hameau Prefontaine | Jul 2019 | N | 96 | OM630668 | OM350266 | OM542394 |
| Cx_spissipes_GF_S97 | Culex (Melanoconion) sp. DJS-2020 | French Guiana | Hameau Prefontaine | Jul 2019 | N | 97 | OM630677 | OM350274 | OM542401 |
| Cx_spissipes_GF_S98 | Culex (Melanoconion) sp. DJS-2020 | French Guiana | Hameau Prefontaine | Jul 2019 | N | 98 | OM630678 | OM350275 | OM542402 |
| Cx_spissipes_GF_S99 | Culex (Melanoconion) sp. DJS-2020 | French Guiana | Hameau Prefontaine | Jul 2019 | N | 99 | OM630679 | OM350276 | OM542403 |
| Li_durhamii_GF_S100 | Limatus durhamii | French Guiana | Hameau Prefontaine | Jul 2019 | N | 100 | OM630695 | OM350291 | OM542419 |
| Li_durhamii_GF_S101 | Limatus durhamii | French Guiana | Hameau Prefontaine | Jul 2019 | N | 101 | OM630696 | OM350292 | OM542420 |
| Li_durhamii_GF_S102 | Limatus durhamii | French Guiana | Hameau Prefontaine | Jul 2019 | N | 102 | OM630697 | OM350293 | OM542421 |
| Ma_sp.4_GF_S103 | Mansonia sp.4 | French Guiana | Hameau Prefontaine | Jan 2020 | N | 103 | OM630701 | OM350298 | OM542425 |
| Ma_sp.4_GF_S104 | Mansonia sp.4 Mansonia | French Guiana | Hameau Prefontaine | Jan 2020 | N | 104 | OM630702 | OM350299 | OM542426 |
| Ma_titillans_GF_S105 | (Mansonia) titillans Culex | French Guiana | Hameau Prefontaine | Jan 2020 | N | 105 | OM630703 | OM350300 | OM542427 |
| Cx_pedroi_GF_S106 | (Melanoconion) pedroi Culex | French Guiana | Hameau Prefontaine | Nov 2020 | N | 106 | OM630663 | OM350261 | OM542389 |
| Cx_pedroi_GF_S107 | (Melanoconion) pedroi Culex | French Guiana | Hameau Prefontaine | Nov 2020 | N | 107 | OM630664 | OM350262 | OM542390 |
| Cx_pedroi_GF_S108 | (Melanoconion) pedroi Psorophora | French Guiana | Hameau Prefontaine | Nov 2020 | N | 108 | OM630665 | OM350263 | OM542391 |
| Ps_ferox_GF_S109 | ferox Psorophora | French Guiana | Iracoubo | 2009 | N | 109 | OM630711 | OM350309 | OM542436 |
| Ps_ferox_GF_S110 | ferox Psorophora | French Guiana | Iracoubo | 2009 | N | 110 | OM630712 | OM350310 | OM542437 |
| Ps_ferox_GF_S111 | ferox Psorophora | French Guiana | Iracoubo | 2009 | N | 111 | OM630713 | OM350324 | OM542452 |
| Ur_geometrica_GF_S112 | Uranotaenia (Uranotaenia) geometrica | French Guiana | | 2010 | N | 112 | OM630714 | OM350311 | OM542453 |
| Ur_geometrica_GF_S113 | Uranotaenia | French | | 2010 | N | 113 | N/A | OM350312 | OM542454 |

| | | | | | | | | | | |
|------------------------------|--|------------------|--------------|----------|---|-----|----------|----------|-----------------------|--|
| | (Uranotaenia) geometrica | Guiana | | | | | | | | |
| Ur_geometrica_GF_S114 | Uranotaenia (Uranotaenia) geometrica | French Guiana | | 2010 | N | 114 | OM630715 | OM350313 | OM542438, OM542455 | |
| Cx_tritaeniorhynchus_MG_S115 | Culex (Culex) tritaeniorhynchus | Madagascar | Ambato Boeny | Feb 2019 | N | 115 | OM630685 | OM350321 | OM542448 | |

727

728 **Appendix 2—table 1.** COI sequence BLAST analyses summary (XLSX).

| Sequence ID | Sequence length | Morphological identification | BLASTn top hit species | BLASTn top hit accession | Query coverage | E-value | % identity | Comments |
|------------------------|-----------------|------------------------------|--|----------------------------|----------------|---------|------------|---|
| Ae_albopictus_KH_S1 | 699 | Aedes albopictus | Aedes albopictus | MK714006.1 | 99% | 0.0 | 99.71% | |
| Ae_albopictus_KH_S2 | 695 | Aedes albopictus | Aedes albopictus | MK714006.1 | 100% | 0.0 | 99.71% | |
| Ae_albopictus_KH_S3 | 695 | Aedes albopictus | Aedes albopictus | MK714006.1 | 100% | 0.0 | 99.71% | |
| An_baezai_KH_S4 | 658 | Anopheles baezai | Anopheles darlingi | MF381626.1 | 100% | 0.0 | 92.71% | An baezai not found in GenBank databases |
| An_baezai_KH_S5 | 670 | Anopheles baezai | Anopheles darlingi | MF381626.1 | 99% | 0.0 | 92.81% | An baezai not found in GenBank databases |
| An_baezai_KH_S6 | 659 | Anopheles baezai | Anopheles darlingi | MF381626.1 | 100% | 0.0 | 92.72% | An baezai not found in GenBank databases |
| Cx_pseudovishnui_KH_S7 | 660 | Culex vishnui | Culex pseudovishnui | MW321882.1 | 98% | 0.0 | 98.92% | 95% similarity to Cx vishnui, 94% similarity with Cx tritaeniorhynchus |
| Cx_pseudovishnui_KH_S8 | 659 | Culex vishnui | Culex pseudovishnui | MW321882.1 | 98% | 0.0 | 99.38% | 95% similarity to Cx vishnui, 94% similarity with Cx tritaeniorhynchus |
| Cx_pseudovishnui_KH_S9 | 659 | Culex vishnui | Culex pseudovishnui | MW321882.1 | 98% | 0.0 | 98.92% | 95% similarity to Cx vishnui, 94% similarity with Cx tritaeniorhynchus |
| Ma_indiana_KH_S10 | 660 | Mansonia indiana | Mansonia indiana | MK637632.1 | 98.00% | 0.0 | 99.54% | |
| Ma_uniformis_KH_S11 | 686 | Mansonia indiana | Mansonia uniformis | MK757484.1 | 99% | 0.0 | 99.71% | 89.99% from Ma indiana MK637632.1 |
| Ma_indiana_KH_S12 | 693 | Mansonia indiana | Mansonia indiana | MK637632.1 | 97% | 0.0 | 99.41% | |
| Cx_sp.1_KH_S13 | 687 | Culex quinquefasciatus | Culex (Lophoceraomyia) sp. 5 HY-2020 | MW321904.1 | 98% | 0.0 | 94.39% | 90% from Cx quinquefasciatus GU188856.2 |
| Cx_orientalis_KH_S14 | 662 | Culex | Culex orientalis | MW228488.1 | 97% | 0.0 | 98.29% | |

| | | quinquefasciatus | | | | | | |
|-----------------------------|-----|-------------------------|-------------------------|----------------------------|---------|-----|---------|---|
| Ma_uniformis_KH_S15 | 658 | Mansonia uniformis | Mansonia uniformis | MK757484.1 | 100.00% | 0.0 | 99.54% | |
| Ma_uniformis_KH_S16 | 654 | Mansonia uniformis | Mansonia uniformis | MK757484.1 | 100.00% | 0.0 | 99.39% | |
| Ma_uniformis_KH_S17 | 657 | Mansonia uniformis | Mansonia uniformis | MK757484.1 | 99.00% | 0.0 | 99.54% | |
| Cx_bitaeniorhynchus_KH_S18 | 658 | Culex bitaeniorhynchus | Culex bitaeniorhynchus | HQ398898.1 | 97.00% | 0.0 | 99.69% | |
| Cx_bitaeniorhynchus_KH_S19 | 650 | Culex bitaeniorhynchus | Culex bitaeniorhynchus | HQ398898.1 | 98.00% | 0.0 | 99.84% | |
| Cx_bitaeniorhynchus_KH_S20 | 652 | Culex bitaeniorhynchus | Culex bitaeniorhynchus | HQ398898.1 | 98.00% | 0.0 | 99.38% | |
| Cx_tritaeniorhynchus_KH_S21 | 695 | Culex tritaeniorhynchus | Culex tritaeniorhynchus | MH374857.1 | 100% | 0.0 | 99.57% | 99.69% to Cx tritaeniorhynchus MF179213.1 |
| Cx_tritaeniorhynchus_KH_S22 | 690 | Culex tritaeniorhynchus | Culex tritaeniorhynchus | MT876103.1 | 100% | 0.0 | 99.57% | |
| Cx_tritaeniorhynchus_KH_S23 | 663 | Culex tritaeniorhynchus | Culex tritaeniorhynchus | MT876103.1 | 99% | 0.0 | 98.79% | |
| Ae_aegypti_CF_S24 | 689 | Aedes aegypti | Aedes aegypti | MN299016.1 | 100% | 0.0 | 99.56% | |
| Ae_aegypti_CF_S25 | 660 | Aedes aegypti | Aedes aegypti | MN299024.1 | 100.00% | 0.0 | 99.70% | |
| Ae_aegypti_CF_S26 | 660 | Aedes aegypti | Aedes aegypti | MN299024.1 | 100.00% | 0.0 | 99.70% | |
| Ae_simpsoni_CF_S27 | 644 | Aedes opok | Aedes simpsoni | LC473669.1 | 97.00% | 0.0 | 97.77% | Ae opok not found in GenBank, sequence is 90% and 89% away from Ae luteocephalus and Ae africanus, sister species of Ae opok. |
| Ae_simpsoni_CF_S28 | 649 | Aedes opok | Aedes simpsoni | MN552302.1 | 99.00% | 0.0 | 100.00% | Ae opok not found in GenBank, sequence is 90% and 89% away from Ae luteocephalus and Ae africanus, sister species of Ae opok. |
| Ae_simpsoni_CF_S29 | 627 | Aedes opok | Aedes simpsoni | MN552302.1 | 98.00% | 0.0 | 98.87% | Ae opok not found in GenBank, sequence is 90% and 89% away from Ae luteocephalus and Ae africanus, sister species of Ae opok. |
| Ae_vittatus_CF_S30 | 623 | Aedes vittatus | Aedes vittatus | MN552298.1 | 100.00% | 0.0 | 99.84% | |
| Ae_vittatus_CF_S31 | 622 | Aedes vittatus | Aedes vittatus | MN552298.1 | 100.00% | 0.0 | 99.68% | |

| | | | | | | | | |
|-----------------------|-----|------------------------|---------------------------|----------------------------|---------|-----|---------|-----------------------------|
| Ae_vittatus_CF_S32 | 621 | Aedes vittatus | Aedes vittatus | MN552298.1 | 100.00% | 0.0 | 99.68% | |
| Ma_sp.1_CF_S33 | - | Mansonia africana | - | - | - | - | - | No COI obtained |
| Ma_sp.2_CF_S34 | - | Mansonia africana | - | - | - | - | - | No COI obtained |
| Ma_sp.3_CF_S35 | - | Mansonia africana | - | - | - | - | - | No COI obtained |
| Ae_albopictus_CF_S36 | 627 | Aedes albopictus | Aedes albopictus | MK995332.1 | 100.00% | 0.0 | 99.84% | |
| Ae_albopictus_CF_S37 | 621 | Aedes albopictus | Aedes albopictus | MK995332.1 | 100.00% | 0.0 | 100.00% | |
| Ae_albopictus_CF_S38 | 621 | Aedes albopictus | Aedes albopictus | MK995332.1 | 100.00% | 0.0 | 100.00% | |
| An_coustani_CF_S39 | 621 | Anopheles coustani | Anopheles coustani | MK585968.1 | 100.00% | 0.0 | 99.84% | |
| An_coustani_CF_S40 | 621 | Anopheles coustani | Anopheles coustani | MK585959.1 | 100.00% | 0.0 | 99.03% | |
| An_coustani_CF_S41 | 699 | Anopheles coustani | Anopheles coustani | MK585968.1 | 94.00% | 0.0 | 99.70% | |
| An_funestus_CF_S42 | 696 | Anopheles funestus | Anopheles funestus | MK300231.1 | 100.00% | 0.0 | 99.71% | |
| An_funestus_CF_S43 | 660 | Anopheles funestus | Anopheles funestus | MT375215.1 | 100.00% | 0.0 | 99.85% | |
| An_funestus_CF_S44 | 658 | Anopheles funestus | Anopheles funestus | MT375215.1 | 100.00% | 0.0 | 99.70% | |
| An_gambiae_CF_S45 | 660 | Anopheles gambiae | Anopheles gambiae | MG930895.1 | 86.00% | 0.0 | 99.79% | |
| An_gambiae_CF_S46 | 659 | Anopheles gambiae | Anopheles gambiae | MT375223.1 | 89.00% | 0.0 | 100.00% | |
| An_gambiae_CF_S47 | - | Anopheles gambiae | - | - | - | - | - | No COI obtained |
| Cx_sp.2_CF_S48 | 653 | Culex quinquefasciatus | Culex corniger | KM593015.1 | 100.00% | 0.0 | 94.95% | 94% to all other Cx species |
| Cx_sp.2_CF_S49 | 660 | Culex quinquefasciatus | Culex nigripalpus | KM593058.1 | 99.00% | 0.0 | 94.65% | 94% to all other Cx species |
| Cx_sp.2_CF_S50 | 658 | Culex quinquefasciatus | Culex bidens | MH931446.1 | 100.00% | 0.0 | 94.68% | 94% to all other Cx species |
| Ma_uniformis_CF_S51 | - | Mansonia uniformis | - | - | - | - | - | No COI obtained |
| Cx_duttoni_CF_S52 | 621 | Mansonia uniformis | Culex duttoni | LC473629.1 | 100.00% | 0.0 | 99.68% | |
| Er_intermedius_CF_S54 | 620 | Eretmapodites sp. | Eretmapodites intermedius | MN552305.1 | 100.00% | 0.0 | 99.52% | |
| Er_intermedius_CF_S55 | 621 | Eretmapodites sp. | Eretmapodites intermedius | MN552305.1 | 100.00% | 0.0 | 99.68% | |

| | | | | | | | | |
|----------------------------|-----|------------------------|---------------------------|-----------------------------|---------|-----|---------|---|
| Er_intermedius_CF_S56 | 621 | Eretmapodites sp. | Eretmapodites intermedius | MN552305.1 | 100.00% | 0.0 | 99.68% | |
| Cx_antennatus_MG_S57 | 621 | Culex antennatus | Culex antennatus | LC473659.1 | 100.00% | 0.0 | 100.00% | |
| Cx_antennatus_MG_S58 | 621 | Culex antennatus | Culex antennatus | LC473659.1 | 100.00% | 0.0 | 100.00% | |
| Cx_antennatus_MG_S59 | 621 | Culex antennatus | Culex antennatus | LC473659.1 | 100.00% | 0.0 | 100.00% | |
| Cx_perexiguus_MG_S60 | 621 | Culex decens | Culex perexiguus | LC473634.1 | 100.00% | 0.0 | 99.84% | |
| Cx_sp.3_MG_S61 | 685 | Culex decens | Unknown Culex species | KU380436.1 | 96.00% | 0.0 | 96.05% | |
| Cx_sp.4_MG_S62 | 687 | Culex decens | Unknown Culex species | MT993494.1 | 99.00% | 0.0 | 95.63% | |
| Cx_sp.3_MG_S63 | 687 | Culex univittatus | Unknown Culex species | KU380436.1 | 95.00% | 0.0 | 96.50% | |
| Mi_sp.1_MG_S64 | 694 | Culex univittatus | Mimomyia mimomyiaformis | LC473719.1 | 94.00% | 0.0 | 92.55% | Unknown Mimomyia species |
| Cx_sp.3_MG_S65 | 691 | Culex univittatus | Unknown Culex species | KU380436.1 | 95.00% | 0.0 | 96.66% | |
| An_coustani_MG_S66 | 669 | Anopheles coustani | Anopheles coustani | NC_050693.1 | 99.00% | 0.0 | 99.40% | |
| An_coustani_MG_S67 | 659 | Anopheles coustani | Anopheles coustani | NC_050693.1 | 99.00% | 0.0 | 99.08% | |
| An_squamosus_MG_S68 | 653 | Anopheles coustani | Anopheles squamosus | MK776741.1 | 100.00% | 0.0 | 100.00% | |
| An_funestus_MG_S69 | 654 | Anopheles funestus | Anopheles funestus | MT375215.1 | 100.00% | 0.0 | 99.85% | |
| An_funestus_MG_S70 | 654 | Anopheles funestus | Anopheles funestus | MG742199.1 | 100.00% | 0.0 | 99.69% | |
| An_gambiae_MG_S71 | 654 | Anopheles gambiae | Anopheles gambiae | MT375222.1 | 100.00% | 0.0 | 99.85% | |
| An_gambiae_MG_S72 | 654 | Anopheles gambiae | Anopheles gambiae | MT375222.1 | 100.00% | 0.0 | 99.85% | |
| An_gambiae_MG_S73 | 622 | Anopheles gambiae | Anopheles gambiae | MT375222.1 | 100.00% | 0.0 | 100.00% | |
| Cx_quinquefasciatus_MG_S74 | 654 | Culex quinquefasciatus | Culex pipiens | MT199095.1 | 100.00% | 0.0 | 100.00% | 99.85% to Cx quinquefasciatus |
| Cx_quinquefasciatus_MG_S75 | 647 | Culex quinquefasciatus | Culex quinquefasciatus | MH423504.1 | 100.00% | 0.0 | 98.15% | Also 98% to Cx pipiens Same SNPs to Cx pipiens MH374861.1 |
| Cx_perexiguus_MG_S76 | 621 | Culex quinquefasciatus | Culex perexiguus | LC473634.1 | 100.00% | 0.0 | 99.52% | |
| Ma_uniformis_MG_S77 | 621 | Mansonia uniformis | Mansonia uniformis | KU187165.1 | 100.00% | 0.0 | 100.00% | |

| | | | | | | | | |
|-------------------------|-----|------------------------------|------------------------------|----------------------------|---------|-----|-----------|---|
| Ma_uniformis_MG_S78 | 621 | Mansonia uniformis | Mansonia uniformis | KU187165.1 | 100.00% | 0.0 | 100.00% | |
| Ma_uniformis_MG_S79 | 626 | Mansonia uniformis | Mansonia uniformis | KU187157.1 | 100.00% | 0.0 | 99.68% | |
| Cx_poecilipes_MG_S82 | 689 | Culex bitaeniorhynchus | Culex poecilipes | LC473618.1 | 95.00% | 0.0 | 99.70% | |
| Mi_mediolineata_MG_S83 | 694 | Culex tritaeniorhynchus | Mimomyia mediolineata | LC473723.1 | 94.00% | 0.0 | 99.39% | |
| Cx_neavei_MG_S84 | 671 | Culex tritaeniorhynchus | Culex neavei | LC473635.1 | 98.00% | 0.0 | 99.85% | |
| Ae_scapularis_GF_S85 | 659 | Aedes scapularis | Aedes scapularis | MN997484.1 | 97.00% | 0.0 | 98.76% | |
| Ae_scapularis_GF_S86 | 658 | Aedes scapularis | Aedes scapularis | MF172265.1 | 97.00% | 0.0 | 99.38% | |
| Ae_scapularis_GF_S87 | 654 | Aedes scapularis | Aedes scapularis | MF172265.1 | 98.00% | 0.0 | 99.22% | |
| Ae_serratus_GF_S88 | 660 | Aedes serratus | Aedes serratus | MF172269.1 | 97.00% | 0.0 | 98.91% | |
| Ae_serratus_GF_S89 | 660 | Aedes serratus | Aedes serratus | MF172268.1 | 97.00% | 0.0 | 99.22% | |
| Ae_serratus_GF_S90 | 654 | Aedes serratus | Aedes serratus | MF172268.1 | 98.00% | 0.0 | 99.07% | |
| Cq_venezuelensis_GF_S91 | 658 | Coquillettidia venezuelensis | Coquillettidia venezuelensis | MN997703.1 | 97.00% | 0.0 | 97.98% | |
| Cq_venezuelensis_GF_S92 | 621 | Coquillettidia venezuelensis | Coquillettidia venezuelensis | MN997703.1 | 100.00% | 0.0 | 98.07% | |
| Cq_venezuelensis_GF_S93 | 621 | Coquillettidia venezuelensis | Coquillettidia venezuelensis | MN997703.1 | 100.00% | 0.0 | 97.75% | |
| Cx_portesi_GF_S94 | 653 | Culex portesi | Culex portesi | in-house reference library | | | 98.5-100% | reference sequence provided by Amandine Guidez, IP Guyane |
| Cx_portesi_GF_S95 | 693 | Culex portesi | Culex portesi | in-house reference library | | | 98.5-100% | reference sequence provided by Amandine Guidez, IP Guyane |
| Cx_portesi_GF_S96 | 687 | Culex portesi | Culex portesi | in-house reference library | | | 98.5-100% | reference sequence provided by Amandine Guidez, IP Guyane |
| Cx_spissipes_GF_S97 | 672 | Culex spissipes | Culex spissipes | in-house reference library | | | 98.5-100% | reference sequence provided by Amandine Guidez, IP Guyane |
| Cx_spissipes_GF_S98 | 663 | Culex spissipes | Culex spissipes | in-house reference library | | | 98.5-100% | reference sequence provided by Amandine Guidez, IP Guyane |
| Cx_spissipes_GF_S99 | 660 | Culex spissipes | Culex spissipes | in-house reference library | | | 98.5-100% | reference sequence provided by Amandine Guidez, IP Guyane |
| Li_durhamii_GF_S100 | 653 | Limatus durhamii | Limatus durhamii | MF172330.1 | 98.00% | 0.0 | 99.84% | |

| | | | | | | | | |
|------------------------------|-----|-------------------------|-------------------------|-----------------------------|---------|-----|---------|--------------------------------------|
| Li_durhamii_GF_S101 | 621 | Limatus durhamii | Limatus durhamii | MF172330.1 | 100.00% | 0.0 | 100.00% | |
| Li_durhamii_GF_S102 | 699 | Limatus durhamii | Limatus durhamii | MF172330.1 | 94.00% | 0.0 | 100.00% | |
| Ma_sp.4_GF_S103 | 621 | Mansonia titillans | Mansonia sp. | MT329066.1 | 100.00% | 0.0 | 99.84% | 87.12% to Ma titillans MN968244.1 |
| Ma_sp.4_GF_S104 | 695 | Mansonia titillans | Mansonia sp. | MT329066.1 | 95.00% | 0.0 | 99.85% | 87.39% to Ma titillans MN968244.1 |
| Ma_titillans_GF_S105 | 669 | Mansonia titillans | Mansonia titillans | MN968244.1 | 98.00% | 0.0 | 99.70% | |
| Cx_pedroi_GF_S106 | 653 | Culex pedroi | Culex pedroi | KX779887.1 | 98.00% | 0.0 | 98.60% | |
| Cx_pedroi_GF_S107 | 661 | Culex pedroi | Culex pedroi | KX779887.1 | 97.00% | 0.0 | 98.76% | |
| Cx_pedroi_GF_S108 | 621 | Culex pedroi | Culex pedroi | KX779887.1 | 99.00% | 0.0 | 98.87% | |
| Ps_ferox_GF_S109 | 633 | Psorophora ferox | Psorophora ferox | MF172349.1 | 100.00% | 0.0 | 99.68% | |
| Ps_ferox_GF_S110 | 621 | Psorophora ferox | Psorophora ferox | MF172349.1 | 100.00% | 0.0 | 99.68% | |
| Ps_ferox_GF_S111 | 621 | Psorophora ferox | Psorophora ferox | MF172347.1 | 99.00% | 0.0 | 99.51% | |
| Ur_geometrica_GF_S112 | 621 | Uranotaenia geometrica | Uranotaenia geometrica | NC_044662.1 | 100.00% | 0.0 | 100.00% | |
| Ur_geometrica_GF_S113 | - | Uranotaenia geometrica | - | - | - | - | - | No COI obtained |
| Ur_geometrica_GF_S114 | 621 | Uranotaenia geometrica | Uranotaenia geometrica | NC_044662.1 | 100.00% | 0.0 | 100.00% | |
| Cx_tritaeniorhynchus_MG_S115 | 653 | Culex tritaeniorhynchus | Culex tritaeniorhynchus | MK861440.1 | 100.00% | 0.0 | 98.77% | |

729 **SOURCE DATA FILES**

730 **Figure 3—source data 1.** Multiple sequence alignment of 169 28S rRNA sequences from this study
731 and from GenBank (FASTA).

732 **Figure 4—source data 1.** Multiple sequence alignment of 122 28S rRNA sequences, including two
733 sequences from *Horreolanus sp.* (FASTA).

734 **Figure 4—source data 2.** Multiple sequence alignment of 114 18S rRNA sequences, including two
735 sequences from *Horreolanus sp.* (FASTA).

736 **Figure 5—source data 1.** Multiple sequence alignment of 106 COI sequences (FASTA).

737 **REFERENCES**

738 Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment
739 search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-](https://doi.org/10.1016/S0022-2836(05)80360-2)
740 2836(05)80360-2

741 Arctander, P. (1995). Comparison of a mitochondrial gene and a corresponding nuclear pseudogene.
742 *Proceedings of the Royal Society B: Biological Sciences*, 262(1363), 13–19.
743 <https://doi.org/10.1098/rspb.1995.0170>

744 Arunachalam, N., Philip Samuel, P., Hiriyan, J., Thenmozhi, V., & Gajanana, A. (2004). Japanese
745 encephalitis in Kerala, South India: Can *Mansonia* (Diptera: Culicidae) play a supplemental role
746 in transmission? *Journal of Medical Entomology*, 41(3), 456–461. [https://doi.org/10.1603/0022-](https://doi.org/10.1603/0022-2585-41.3.456)
747 2585-41.3.456

748 Aspen, S., & Savage, H. M. (2003). Polymerase chain reaction assay identifies North American
749 members of the *Culex pipiens* complex based on nucleotide sequence differences in the
750 acetylcholinesterase gene *Ace2*. *Journal of the American Mosquito Control Association*, 19(4),
751 323–328.

752 Auerswald, H., Maquart, P. O., Chevalier, V., & Boyer, S. (2021). Mosquito vector competence for
753 japanese encephalitis virus. *Viruses*, 13(6), 1154. <https://doi.org/10.3390/v13061154>

754 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M.,
755 Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G.,
756 Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its
757 applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455–477.

758 <https://doi.org/10.1089/cmb.2012.0021>

759 Barrio-Nuevo, K. M., Cunha, M. S., Luchs, A., Fernandes, A., Rocco, I. M., Mucci, L. F., DE Souza, R.
760 P., Medeiros-Sousa, A. R., Ceretti-Junior, W., & Marrelli, M. T. (2020). Detection of Zika and
761 dengue viruses in wildcaught mosquitoes collected during field surveillance in an environmental
762 protection area in São Paulo, Brazil. *PLoS ONE*, *15*(10), e0227239.
763 <https://doi.org/10.1371/journal.pone.0227239>

764 Batovska, J., Cogan, N. O. I., Lynch, S. E., & Blacket, M. J. (2017). Using next-generation sequencing
765 for DNA barcoding: Capturing allelic variation in ITS2. *G3: Genes, Genomes, Genetics*, *7*(1),
766 19–29. <https://doi.org/10.1534/G3.116.036145/-/DC1>

767 Beebe, N. W. (2018). DNA barcoding mosquitoes: Advice for potential prospectors. *Parasitology*,
768 *145*(5), 622–633. <https://doi.org/10.1017/S0031182018000343>

769 Behura, S. K. (2006). Molecular marker systems in insects: current trends and future avenues.
770 *Molecular Ecology*, *15*(11), 3087–3113. <https://doi.org/10.1111/J.1365-294X.2006.03014.X>

771 Belda, E., Nanfack-Minkeu, F., Eiglmeier, K., Carissimo, G., Holm, I., Diallo, M., Diallo, D., Vantaux,
772 A., Kim, S., Sharakhov, I. V., & Vernick, K. D. (2019). De novo profiling of RNA viruses in
773 Anopheles malaria vector mosquitoes from forest ecological zones in Senegal and Cambodia.
774 *BMC Genomics*, *20*(1), 664. <https://doi.org/10.1186/s12864-019-6034-1>

775 Bhattacharya, S., & Basu, P. (2016). The Southern House Mosquito, *Culex quinquefasciatus*: profile
776 of a smart vector. *Journal of Entomology and Zoology Studies JEZS*, *4*(2), 73–81.

777 Bishop-Lilly, K. A., Turell, M. J., Willner, K. M., Butani, A., Nolan, N. M. E., Lentz, S. M., Akmal, A.,
778 Mateczun, A., Brahmhatt, T. N., Sozhamannan, S., Whitehouse, C. A., & Read, T. D. (2010).
779 Arbovirus detection in insect vectors by Rapid, high- throughput pyrosequencing. *PLoS*
780 *Neglected Tropical Diseases*, *4*(11), e878. <https://doi.org/10.1371/journal.pntd.0000878>

781 Brault, A. C., Foy, B. D., Myles, K. M., Kelly, C. L. H., Higgs, S., Weaver, S. C., Olson, K. E., Miller, B.
782 R., & Powers, A. M. (2004). Infection patterns of o'nyong nyong virus in the malaria-transmitting
783 mosquito, *Anopheles gambiae*. *Insect Molecular Biology*, *13*(6), 625–635.
784 <https://doi.org/10.1111/j.0962-1075.2004.00521.x>

785 Cardoso, J. da C., de Almeida, M. A. B., dos Santos, E., da Fonseca, D. F., Sallum, M. A. M., Noll, C.
786 A., Monteiro, H. A. d. O., Cruz, A. C. R., Carvalho, V. L., Pinto, E. V., Castro, F. C., Neto, J. P.
787 N., Segura, M. N. O., & Vasconcelos, P. F. C. (2010). Yellow fever virus in *Haemagogus*

788 leucocelaenus and Aedes serratus mosquitoes, Southern Brazil, 2008. *Emerging Infectious*
789 *Diseases*, 16(12), 1918–1924. <https://doi.org/10.3201/eid1612.100608>

790 Chandler, J. A., Liu, R. M., & Bennett, S. N. (2015). RNA Shotgun Metagenomic Sequencing of
791 Northern California (USA) Mosquitoes Uncovers Viruses, Bacteria, and Fungi. *Frontiers in*
792 *Microbiology*, 6, 185. <https://doi.org/10.3389/fmicb.2015.00185>

793 Cornel, A. J., Mcabee, R. D., Rasgon, J., Stanich, M. A., Scott, T. W., & Coetzee, M. (2003).
794 Differences in Extent of Genetic Introgression between Sympatric Culex pipiens and Culex
795 quinquefasciatus (Diptera: Culicidae) in California and South Africa. *Journal of Medical*
796 *Entomology*, 40(1), 36–51. <https://doi.org/10.1603/0022-2585-40.1.36>

797 Danforth, B. N., Lin, C. P., & Fang, J. (2005). How do insect nuclear ribosomal genes compare to
798 protein-coding genes in phylogenetic utility and nucleotide substitution patterns? *Systematic*
799 *Entomology*, 30(4), 549–562. <https://doi.org/10.1111/J.1a365-3113.2005.00305.X>

800 De Oliveira, C. D., Gonçalves, D. S., Baton, L. A., Shimabukuro, P. H. F., Carvalho, F. D., & Moreira,
801 L. A. (2015). Broader prevalence of Wolbachia in insects including potential human disease
802 vectors. *Bulletin of Entomological Research*, 105(3), 305–315.
803 <https://doi.org/10.1017/S0007485315000085>

804 Desdouits, M., Kamgang, B., Berthet, N., Tricou, V., Ngoagouni, C., Gessain, A., Manuguerra, J. C.,
805 Nakouné, E., & Kazanji, M. (2015). Genetic characterization of Chikungunya virus in the Central
806 African Republic. *Infection, Genetics and Evolution*, 33, 25–31.
807 <https://doi.org/10.1016/J.MEEGID.2015.04.006>

808 Diallo, D., Fall, G., Diagne, C. T., Gaye, A., Ba, Y., Dia, I., Faye, O., & Diallo, M. (2020). Concurrent
809 amplification of Zika, chikungunya, and yellow fever virus in a sylvatic focus of arboviruses in
810 Southeastern Senegal, 2015. *BMC Microbiology*, 20, 181. [https://doi.org/10.1186/s12866-020-](https://doi.org/10.1186/s12866-020-01866-9)
811 [01866-9](https://doi.org/10.1186/s12866-020-01866-9)

812 Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space
813 complexity. *BMC Bioinformatics*, 5, 113. <https://doi.org/10.1186/1471-2105-5-113>

814 Edwards, F. W. (1941). *Mosquitoes of the Ethiopian Region: III. Culicine Adults and Pupae*. Order of
815 the Trustees.

816 Farajollahi, A., Fonseca, D. M., Kramer, L. D., & Marm Kilpatrick, A. (2011). “Bird biting” mosquitoes
817 and human disease: A review of the role of Culex pipiens complex mosquitoes in epidemiology.

818 *Infection, Genetics and Evolution*, 11(7), 1577–1585. doi: 10.1016/j.meegid.2011.08.013

819 Fauver, J. R., Akter, S., Morales, A. I. O., Black, W. C., Rodriguez, A. D., Stenglein, M. D., Ebel, G.
820 D., & Weger-Lucarelli, J. (2019). A reverse-transcription/RNase H based protocol for depletion of
821 mosquito ribosomal RNA facilitates viral intrahost evolution analysis, transcriptomics and
822 pathogen discovery. *Virology*, 528, 181–197. <https://doi.org/10.1016/j.virol.2018.12.020>

823 Foley, D. H., Rueda, L. M., & Wilkerson, R. C. (2007). Insight into Global Mosquito Biogeography from
824 Country Species Records. *Journal of Medical Entomology*, 44(4), 554–567.
825 <https://doi.org/10.1093/JMEDENT/44.4.554>

826 Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of
827 mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular*
828 *Marine Biology and Biotechnology*, 3(5), 294–299.

829 Gale, K., & Crampton, J. (1989). The ribosomal genes of the mosquito, *Aedes aegypti*. *European*
830 *Journal of Biochemistry*, 185(2), 311–317. <https://doi.org/10.1111/j.1432-1033.1989.tb15117.x>

831 Grjebine, A. (1966). *Insectes Diptères Culicidae Anophelinae*. ORSTOM / CNRS.

832 Hajibabaei, M., Singer, G. A. C., & Hickey, D. A. (2006). Benchmarking DNA barcodes: An
833 assessment using available primate sequences. *Genome*, 49(7), 851–854.
834 https://doi.org/10.1139/G06-025/SUPPL_FILE/G06-025B.PDF

835 Halstead, S. B. (2019). Travelling arboviruses: A historical perspective. *Travel Medicine and*
836 *Infectious Disease*, 31, 101471. <https://doi.org/10.1016/J.TMAID.2019.101471>

837 Harbach, R. E. (2007). The Culicidae (Diptera): A review of taxonomy, classification and phylogeny.
838 *Zootaxa*, 1668(1), 591–638. <https://doi.org/10.11646/zootaxa.1668.1.28>

839 Harbach, R. E., Culverwell, C. L., & Kitching, I. J. (2017). Phylogeny of the nominotypical subgenus of
840 *Culex* (Diptera: Culicidae): insights from analyses of anatomical data into interspecific
841 relationships and species groups in an unresolved tree. *Systematics and Biodiversity*, 15(4),
842 296–306. <https://doi.org/10.1080/14772000.2016.1252439>

843 Harbach, R. E., & Kitching, I. J. (2016). The phylogeny of Anophelinae revisited: Inferences about the
844 origin and classification of Anopheles (Diptera: Culicidae). *Zoologica Scripta*, 45(1), 34–47.
845 <https://doi.org/10.1111/zsc.12137>

846 Hayes, C. G., Basit, A., Bagar, S., & Akhter, R. (1980). Vector competence of *Culex tritaeniorhynchus*
847 (Diptera: Culicidae) for West Nile virus. *Journal of Medical Entomology*.

848 <https://doi.org/10.1093/jmedent/17.2.172>

849 Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through
850 DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), 313–321.
851 <https://doi.org/10.1098/rspb.2002.2218>

852 Héraud, J.-M., Andriamandimby, S. F., Olive, M.-M., Guis, H., Miatrana Rasamoelina, V., & Tantely, L.
853 (2022). Arthropod-Borne Viruses of Madagascar. In S. M. Goodman (Ed.), *The New Natural*
854 *History of Madagascar* (pp. 285–291). Princeton University Press.

855 Hoyos-López, R., Soto, S. U., Rúa-Urbe, G., & Gallego-Gómez, J. C. (2015). Molecular identification
856 of saint louis encephalitis virus genotype IV in Colombia. *Memorias Do Instituto Oswaldo Cruz*,
857 110(6), 719–725. <https://doi.org/10.1590/0074-02760280040>

858 Huang, Y., & Ward, R. A. (1981). A Pictorial Key for the Identification of the Mosquitoes Associated
859 with Yellow Fever in Africa. *Mosquito Systematics*.

860 Hurst, G. D. D., & Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population,
861 phylogeographic and phylogenetic studies: The effects of inherited symbionts. *Proceedings of*
862 *the Royal Society B: Biological Sciences*, 272, 1525–1534.
863 <https://doi.org/10.1098/rspb.2005.3056>

864 Jacobi, J. C., & Serie, C. (1972). Prevalence of group B arbovirus infections in French Guiana in
865 1967-69. *Medecine d'Afrique Noire*, 19(3), 225–226.

866 Jupp, P. G., Kemp, A., Grobbelaar, A., Leman, P., Burt, F. J., Alahmed, A. M., Al Mujalli, D., Al
867 Khamees, M., & Swanepoel, R. (2002). The 2000 epidemic of Rift Valley fever in Saudi Arabia:
868 Mosquito vector studies. *Medical and Veterinary Entomology*. [https://doi.org/10.1046/j.1365-](https://doi.org/10.1046/j.1365-2915.2002.00371.x)
869 [2915.2002.00371.x](https://doi.org/10.1046/j.1365-2915.2002.00371.x)

870 Kim, H., Cha, G. W., Jeong, Y. E., Lee, W. G., Chang, K. S., Roh, J. Y., Yang, S. C., Park, M. Y.,
871 Park, C., & Shin, E. H. (2015). Detection of Japanese encephalitis virus genotype V in *Culex*
872 *orientalis* and *Culex pipiens* (Diptera: Culicidae) in Korea. *PLoS ONE*, 10(2), e0116547.
873 <https://doi.org/10.1371/journal.pone.0116547>

874 Kraemer, M. U. G., Reiner, R. C., Brady, O. J., Messina, J. P., Gilbert, M., Pigott, D. M., Yi, D.,
875 Johnson, K., Earl, L., Marczak, L. B., Shirude, S., Davis Weaver, N., Bisanzio, D., Perkins, T. A.,
876 Lai, S., Lu, X., Jones, P., Coelho, G. E., Carvalho, R. G., ... Golding, N. (2019). Past and future
877 spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology*, 4(5),

878 854–863. <https://doi.org/10.1038/s41564-019-0376-y>

879 Kukutla, P., Steritz, M., & Xu, J. (2013). Depletion of ribosomal RNA for mosquito gut metagenomic
880 RNA-seq. *Journal of Visualized Experiments*, 74, 50093. <https://doi.org/10.3791/50093>

881 Kumar, N., Creasy, T., Sun, Y., Flowers, M., Tallon, L. J., & Dunning Hotopp, J. C. (2012). Efficient
882 subtraction of insect rRNA prior to transcriptome analysis of Wolbachia-Drosophila lateral gene
883 transfer. *BMC Research Notes*, 5, 230. <https://doi.org/10.1186/1756-0500-5-230>

884 Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary
885 genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), 1547–
886 1549. <https://doi.org/10.1093/molbev/msy096>

887 Logue, K., Chan, E. R., Phipps, T., Small, S. T., Reimer, L., Henry-Halldin, C., Sattabongkot, J., Siba,
888 P. M., Zimmerman, P. A., & Serre, D. (2013). Mitochondrial genome sequences reveal deep
889 divergences among Anopheles punctulatus sibling species in Papua New Guinea. *Malaria*
890 *Journal*, 12(1), 1–11. <https://doi.org/10.1186/1475-2875-12-64/FIGURES/3>

891 Lorenz, C., Alves, J. M. P., Foster, P. G., Suesdek, L., & Sallum, M. A. M. (2021). Phylogeny and
892 temporal diversification of mosquitoes (Diptera: Culicidae) with an emphasis on the Neotropical
893 fauna. *Systematic Entomology*, 46(4), 798–811. <https://doi.org/10.1111/SYEN.12489>

894 Lutomiah, J., Bast, J., Clark, J., Richardson, J., Yalwala, S., Oullo, D., Mutisya, J., Mulwa, F., Musila,
895 L., Khamadi, S., Schnabel, D., Wurapa, E., & Sang, R. (2013). Abundance, diversity, and
896 distribution of mosquito vectors in selected ecological regions of Kenya: Public health
897 implications. *Journal of Vector Ecology*, 38(1), 134–142. [https://doi.org/10.1111/j.1948-](https://doi.org/10.1111/j.1948-7134.2013.12019.x)
898 [7134.2013.12019.x](https://doi.org/10.1111/j.1948-7134.2013.12019.x)

899 Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N.,
900 Potter, S. C., Finn, R. D., & Lopez, R. (2019). The EMBL-EBI search and sequence analysis
901 tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641.
902 <https://doi.org/10.1093/nar/gkz268>

903 Maquart, P. O., & Boyer, S. (2022). Culex vishnui. *Trends in Parasitology, Vector of the Month*, 491–
904 492. <https://doi.org/10.1016/J.PT.2022.01.003>

905 Maquart, P. O., Sokha, C., & Boyer, S. (2021). Mosquito diversity (Diptera: Culicidae) and medical
906 importance, in a bird sanctuary inside the flooded forest of Prek Toal, Cambodia. *Journal of*
907 *Asia-Pacific Entomology*, 24(4), 1221–1227. <https://doi.org/10.1016/j.aspen.2021.08.001>

908 Mitchell, C. J., Forattini, O. P., & Miller, B. R. (1986). Vector competence experiments with Rocio virus
909 and three mosquito species from the epidemic zone in Brazil. *Revista de Saúde Pública*, *20*(3),
910 171–177. <https://doi.org/10.1590/s0034-89101986000300001>

911 Morlan, J. D., Qu, K., & Sinicropi, D. V. (2012). Selective depletion of rRNA enables whole
912 transcriptome profiling of archival fixed tissue. *PLoS ONE*, *7*(8), e42882.
913 <https://doi.org/10.1371/journal.pone.0042882>

914 Mukwaya, L. G., Kayondo, J. K., Crabtree, M. B., Savage, H. M., Biggerstaff, B. J., & Miller, B. R.
915 (2000). Genetic differentiation in the yellow fever virus vector, *Aedes simpsoni* complex, in
916 Africa: Sequence variation in the ribosomal DNA internal transcribed spacers of anthropophilic
917 and non-anthropophilic populations. *Insect Molecular Biology*, *9*(1), 85–91. doi: 10.1046/j.1365-
918 2583.2000.00161.x

919 Mwangangi, J. M., Muturi, E. J., Muriu, S. M., Nzovu, J., Midega, J. T., & Mbogo, C. (2013). The role
920 of *Anopheles arabiensis* and *Anopheles coustani* in indoor and outdoor malaria transmission in
921 Taveta District, Kenya. *Parasites and Vectors*, *6*, 114. <https://doi.org/10.1186/1756-3305-6-114>

922 Navarro, J. C., & Weaver, S. C. (2004). Molecular phylogeny of the Vomerifer and Pedroi Groups in
923 the spissipes section of the subgenus *Culex* (Melanoconion). *Journal of Medical Entomology*,
924 *41*(4), 575–581. <https://doi.org/10.1603/0022-2585-41.4.575>

925 Nchoutpouen, E., Talipouo, A., Djiappi-Tchamen, B., Djamouko-Djonkam, L., Kopya, E., Ngadjeu, C.
926 S., Doumbe-Belisse, P., Awono-Ambene, P., Kekeunou, S., Wondji, C. S., & Antonio-Nkondjio,
927 C. (2019). *Culex* species diversity, susceptibility to insecticides and role as potential vector of
928 Lymphatic filariasis in the city of Yaoundé, Cameroon. *PLoS Neglected Tropical Diseases*,
929 *13*(4), e0007229. <https://doi.org/10.1371/journal.pntd.0007229>

930 Ndiaye, E. H., Fall, G., Gaye, A., Bob, N. S., Talla, C., Diagne, C. T., Diallo, D., Ba, Y., Dia, I., Kohl,
931 A., Sall, A. A., & Diallo, M. (2016). Vector competence of *Aedes vexans* (Meigen), *Culex*
932 *poicilipes* (Theobald) and *Cx. quinquefasciatus* Say from Senegal for West and East African
933 lineages of Rift Valley fever virus. *Parasites and Vectors*, *9*, 94. [https://doi.org/10.1186/s13071-](https://doi.org/10.1186/s13071-016-1383-y)
934 [016-1383-y](https://doi.org/10.1186/s13071-016-1383-y)

935 Nepomichene, T. N. J. J., Raharimalala, F. N., Andriamandimby, S. F., Ravalohery, J. P., Failloux, A.
936 B., Heraud, J. M., & Boyer, S. (2018). Vector competence of *Culex antennatus* and *Anopheles*
937 *coustani* mosquitoes for Rift Valley fever virus in Madagascar. *Medical and Veterinary*

938 *Entomology*, 32(2), 259–262. <https://doi.org/10.1111/mve.12291>

939 Nikolay, B., Diallo, M., Boye, C. S. B., & Sall, A. A. (2011). Usutu virus in Africa. *Vector-Borne and*
940 *Zoonotic Diseases*, 11(11), 1417–1423. <https://doi.org/10.1089/vbz.2011.0631>

941 Oo, T. T., Kaiser, A., & Becker, N. (2006). Illustrated keys to the anopheline mosquitoes of Myanmar.
942 *Journal of Vector Ecology*, 31(1), 9–16. [https://doi.org/10.3376/1081-](https://doi.org/10.3376/1081-1710(2006)31[9:ikttam]2.0.co;2)
943 [1710\(2006\)31\[9:ikttam\]2.0.co;2](https://doi.org/10.3376/1081-1710(2006)31[9:ikttam]2.0.co;2)

944 Pereira Serra, O., Fernandes Cardoso, B., Maria Ribeiro, A. L., dos Santos, F. A. L., & Dezengrini
945 Shhessarenko, R. (2016). Mayaro virus and dengue virus 1 and 4 natural infection in culicids
946 from Cuiabá, state of Mato Grosso, Brazil. *Memórias Do Instituto Oswaldo Cruz*, 111(1), 20–29.
947 <https://doi.org/10.1590/0074-02760150270>

948 Phelps, W. A., Carlson, A. E., & Lee, M. T. (2021). Optimized design of antisense oligomers for
949 targeted rRNA depletion. *Nucleic Acids Research*, 49(1), e5.
950 <https://doi.org/10.1093/nar/gkaa1072>

951 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O.
952 (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-
953 based tools. *Nucleic Acids Research*, 41(Database issue), D590–D596.
954 <https://doi.org/10.1093/nar/gks1219>

955 Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System: Barcoding.
956 *Molecular Ecology Notes*, 7(3), 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>

957 Ratovonjato, J., Olive, M. M., Tantely, L. M., Andrianaivolambo, L., Tata, E., Razainirina, J.,
958 Jeanmaire, E., Reynes, J. M., & Elissa, N. (2011). Detection, isolation, and genetic
959 characterization of Rift Valley fever virus from anopheles (*Anopheles*) coustani, anopheles
960 (*Anopheles*) squamosus, and culex (*Culex*) antennatus of the haute matsiatra region,
961 Madagascar. *Vector-Borne and Zoonotic Diseases*, 11(6), 753–759.
962 <https://doi.org/10.1089/vbz.2010.0031>

963 Ratsitorahina, M., Harisoa, J., Ratovonjato, J., Biacabe, S., Reynes, J. M., Zeller, H., Raelina, Y.,
964 Talarmin, A., Richard, V., & Soares, J. L. (2008). Outbreak of dengue and chikungunya fevers,
965 Toamasina, Madagascar, 2006. *Emerging Infectious Diseases*, 14(7), 1135–1137.
966 <https://doi.org/10.3201/EID1407.071521>

967 Rattanarithkul, R., Harbach, R. E., Harrison, B. A., Panthusiri, P., & Coleman, R. E. (2007). Illustrated

968 keys to the mosquitoes of Thailand V. Genera Orthopodomyia, Kimia, Malaya, Topomyia,
969 Tripteroides, and Toxorhynchites. *Suppl 1*, 38(Suppl 2), 1–65.

970 Rattanarithikul, R., Harbach, R. E., Harrison, B. A., Panthusiri, P., Coleman, R. E., & Richardson, J. H.
971 (2010). Illustrated keys to the mosquitoes of Thailand. VI. Tribe Aedini. *Southeast Asian Journal*
972 *of Tropical Medicine and Public Health*, 41(Suppl 1), 1–225.

973 Rattanarithikul, R., Harbach, R. E., Harrison, B. A., Panthusiri, P., Jones, J. W., & Coleman, R. E.
974 (2005). Illustrated keys to the mosquitoes of Thailand. II. Genera Culex and Lutzia. *The*
975 *Southeast Asian Journal of Tropical Medicine and Public Health*.

976 Rattanarithikul, R., Harrison, B. A., Harbach, R. E., Panthusiri, P., & Coleman, R. E. (2006). Illustrated
977 keys to the mosquitoes of Thailand IV. Anopheles. *Southeast Asian Journal of Tropical Medicine*
978 *and Public Health*, 37(Suppl 2), 1–26.

979 Rattanarithikul, R., Harrison, B. A., Panthusiri, P., & Coleman, R. E. (2005). Illustrated keys to the
980 mosquitoes of Thailand I. Background; geographic distribution; lists of genera, subgenera, and
981 species; and a key to the genera. *Southeast Asian Journal of Tropical Medicine and Public*
982 *Health*, 36(Suppl 1), 1–80.

983 Rattanarithikul, R., Harrison, B. A., Panthusiri, P., Peyton, E. L., & Coleman, R. E. (2006). Illustrated
984 keys to the mosquitoes of Thailand: III. Genera Aedeomyia, Ficalbia, Mimomyia, Hodgesia,
985 Coquillettidia, Mansonia, and Uranotaenia. *Southeast Asian Journal of Tropical Medicine and*
986 *Public Health*, 37(Suppl 1), 1–10.

987 Rausch, T., Fritz, M. H. Y., Untergasser, A., & Benes, V. (2020). Tracy: Basecalling, alignment,
988 assembly and deconvolution of sanger chromatogram trace files. *BMC Genomics*, 21(1), 230.
989 <https://doi.org/10.1186/s12864-020-6635-8>

990 Rausch, T., Hsi-Yang Fritz, M., Korb, J. O., & Benes, V. (2019). Alfred: Interactive multi-sample
991 BAM alignment statistics, feature counting and feature annotation for long- and short-read
992 sequencing. *Bioinformatics*, 35(14), 2489–2491. <https://doi.org/10.1093/bioinformatics/bty1007>

993 Reidenbach, K. R., Cook, S., Bertone, M. A., Harbach, R. E., Wiegmann, B. M., & Besansky, N. J.
994 (2009). Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae)
995 based on nuclear genes and morphology. *BMC Evolutionary Biology*, 9(1), 1–14.
996 <https://doi.org/10.1186/1471-2148-9-298/FIGURES/4>

997 Romero-Alvarez, D., & Escobar, L. E. (2018). Oropouche fever, an emergent disease from the

998 Americas. *Microbes and Infection*, 20(3), 135–146.

999 <https://doi.org/10.1016/J.MICINF.2017.11.013>

1000 Rueda, L. M. (2004). Pictorial keys for the identification of mosquitoes (Diptera: Culicidae) associated
1001 with Dengue Virus Transmission. *Zootaxa*. <https://doi.org/10.11646/zootaxa.589.1.1>

1002 Ruzzante, L., Reijnders, M. J. M. F., & Waterhouse, R. M. (2019). Of Genes and Genomes: Mosquito
1003 Evolution and Diversity. *Trends in Parasitology*, 35(1), 32–51.
1004 [https://doi.org/10.1016/J.PT.2018.10.003/ATTACHMENT/B9BE6BC5-D73A-4FEF-A654-
1005 CC1374C59925/MMC1.MP4](https://doi.org/10.1016/J.PT.2018.10.003/ATTACHMENT/B9BE6BC5-D73A-4FEF-A654-CC1374C59925/MMC1.MP4)

1006 Sallum, M. A. M., & Forattini, O. P. (1996). Revision of the Spissipes Section of Culex (Melanoconion)
1007 (diptera: Culicidae). *Journal of the American Mosquito Control Association*, 12(3), 517–600.

1008 Saluzzo, J. F., Evidera, T. V., Veas, F., & Gonzalez, J.-P. J. (2018). *Arbovirus Discovery in Central
1009 African Republic (1973-1993): Zika, Bozo, Bouboui, and More*.
1010 <https://www.researchgate.net/publication/321587457>

1011 Sirivanakarn, S. (1982). A review of the Systematics and a Proposed Scheme of Internal
1012 Classification of the New World Subgenus Melanoconion of Culex (Diptera, Culicidae). *Mosquito
1013 Systematics*, 14(4), 265–333.

1014 Stevenson, J. C., Simubali, L., Mbambara, S., Musonda, M., Mweetwa, S., Mudenda, T., Pringle, J.
1015 C., Jones, C. M., & Norris, D. E. (2016). Detection of plasmodium falciparum infection in
1016 anopheles squamosus (diptera: Culicidae) in an area targeted for malaria elimination, Southern
1017 Zambia. *Journal of Medical Entomology*, 53(6), 1482–1487. <https://doi.org/10.1093/jme/tjw091>

1018 Sun, L., Li, T. J., Fu, W. B., Yan, Z. T., Si, F. L., Zhang, Y. J., Mao, Q. M., Demari-Silva, B., & Chen,
1019 B. (2019). The complete mt genomes of Lutzia halifaxia, Lt. fuscanus and Culex pallidothorax
1020 (Diptera: Culicidae) and comparative analysis of 16 Culex and Lutzia mt genome sequences.
1021 *Parasites and Vectors*, 12, 368. <https://doi.org/10.1186/s13071-019-3625-2>

1022 Tabue, R. N., Awono-Ambene, P., Etang, J., Atangana, J., Antonio-Nkondjio, C., Toto, J. C.,
1023 Patchoke, S., Leke, R. G. F., Fondjo, E., Mnzava, A. P., Knox, T. B., Tougordi, A., Donnelly, M.
1024 J., & Bigoga, J. D. (2017). Role of Anopheles (Cellia) rufipes (Gough, 1910) and other local
1025 anophelines in human malaria transmission in the northern savannah of Cameroon: a cross-
1026 sectional survey. *Parasites and Vectors*, 10(1), 1–11. [https://doi.org/10.1186/S13071-016-1933-
1027 3/FIGURES/6](https://doi.org/10.1186/S13071-016-1933-3/FIGURES/6)

1028 Takhampunya, R., Kim, H. C., Tippayachai, B., Kengluetcha, A., Klein, T. A., Lee, W. J., Grieco, J., &
1029 Evans, B. P. (2011). Emergence of Japanese encephalitis virus genotype v in the Republic of
1030 Korea. *Virology Journal*, 8, 449. <https://doi.org/10.1186/1743-422X-8-449>

1031 Talaga, S., Duchemin, J. B., Girod, R., & Dusfour, I. (2021). The Culex Mosquitoes (Diptera:
1032 Culicidae) of French Guiana: A Comprehensive Review With the Description of Three New
1033 Species. *Journal of Medical Entomology*, 58(1), 182–221. <https://doi.org/10.1093/JME/TJAA205>

1034 Thongsripong, P., Chandler, J. A., Kittayapong, P., Wilcox, B. A., Kapan, D. D., & Bennett, S. N.
1035 (2021). Metagenomic shotgun sequencing reveals host species as an important driver of virome
1036 composition in mosquitoes. *Scientific Reports*, 11(1), 8448. [https://doi.org/10.1038/s41598-021-](https://doi.org/10.1038/s41598-021-87122-0)
1037 87122-0

1038 Torres-Gutierrez, C., Bergo, E. S., Emerson, K. J., de Oliveira, T. M. P., Greni, S., & Sallum, M. A. M.
1039 (2016). Mitochondrial COI gene as a tool in the taxonomy of mosquitoes Culex subgenus
1040 Melanoconion. *Acta Tropica*, 164, 137–149. <https://doi.org/10.1016/j.actatropica.2016.09.007>

1041 Torres-Gutierrez, C., De Oliveira, T. M. P., Emerson, K. J., Bergo, E. S., & Sallum, M. A. M. (2018).
1042 Molecular phylogeny of Culex subgenus Melanoconion (Diptera: Culicidae) based on nuclear
1043 and mitochondrial protein-coding genes. *Royal Society Open Science*, 5, 171900.
1044 <https://doi.org/10.1098/rsos.171900>

1045 Travassos Da Rosa, J. F., De Souza, W. M., De Paula Pinheiro, F., Figueiredo, M. L., Cardoso, J. F.,
1046 Acrani, G. O., & Teixeira Nunes, M. R. (2017). Oropouche Virus: Clinical, Epidemiological, and
1047 Molecular Aspects of a Neglected Orthobunyavirus. *The American Journal of Tropical Medicine*
1048 *and Hygiene*, 96(5), 1019. <https://doi.org/10.4269/AJTMH.16-0672>

1049 Turell, M. J., O'guinn, M. L., Dohm, D., Zyzak, M., Watts, D., Fernandez, R., Calampa, C., Klein, T. A.,
1050 & Jones, J. W. (2008). Susceptibility of Peruvian Mosquitoes to Eastern Equine Encephalitis
1051 Virus. *Journal of Medical Entomology*, 45(4), 720–725.
1052 <https://doi.org/10.1093/JMEDENT/45.4.720>

1053 Turell, Michael J. (1999). Vector competence of three Venezuelan mosquitoes (Diptera: Culicidae) for
1054 an epizootic IC strain of Venezuelan equine encephalitis virus. *Journal of Medical Entomology*,
1055 36(4), 407–409. <https://doi.org/10.1093/jmedent/36.4.407>

1056 Ughasi, J., Bekard, H. E., Coulibaly, M., Adabie-Gomez, D., Gyapong, J., Appawu, M., Wilson, M. D.,
1057 & Boakye, D. A. (2012). *Mansonia africana* and *Mansonia uniformis* are Vectors in the

1058 transmission of *Wuchereria bancrofti* lymphatic filariasis in Ghana. *Parasites and Vectors*, 5(1),
1059 1–5. <https://doi.org/10.1186/1756-3305-5-89>

1060 Valentine, M. J., Murdock, C. C., & Kelly, P. J. (2019). Sylvatic cycles of arboviruses in non-human
1061 primates. *Parasites and Vectors*, 12(1), 1–18. [https://doi.org/10.1186/S13071-019-3732-](https://doi.org/10.1186/S13071-019-3732-0/TABLES/4)
1062 [0/TABLES/4](https://doi.org/10.1186/S13071-019-3732-0/TABLES/4)

1063 Vasconcelos, P. F. C., Costa, Z. G., Travassos da Rosa, E. S., Luna, E., Rodrigues, S. G., Barros, V.
1064 L. R. S., Dias, J. P., Monteiro, H. A. O., Oliva, O. F. P., Vasconcelos, H. B., Oliveira, R. C.,
1065 Sousa, M. R. S., Barbosa Da Silva, J., Cruz, A. C. R., Martins, E. C., & Travassos Da Rosa, J.
1066 F. S. (2001). Epidemic of jungle yellow fever in Brazil, 2000: Implications of climatic alterations in
1067 disease spread. *Journal of Medical Virology*, 65(3), 598–604.
1068 <https://doi.org/10.1002/jmv.2078.abs>

1069 Vázquez González, A., Ruiz, S., Herrero, L., Moreno, J., Molero, F., Magallanes, A., Sánchez-Seco,
1070 M. P., Figuerola, J., & Tenorio, A. (2011). West Nile and Usutu viruses in mosquitoes in Spain,
1071 2008-2009. *American Journal of Tropical Medicine and Hygiene*, 85(1), 178–181.
1072 <https://doi.org/10.4269/ajtmh.2011.11-0042>

1073 Vezenegho, S. B., Issaly, J., Carinci, R., Gaborit, P., Girod, R., Dusfour, I., & Briolant, S. (2022).
1074 Discrimination of 15 Amazonian Anopheline Mosquito Species by Polymerase Chain Reaction—
1075 Restriction Fragment Length Polymorphism. *Journal of Medical Entomology*, 59(3), 1060–1064.
1076 <https://doi.org/10.1093/JME/TJAC008>

1077 Weaver, S. C., Ferro, C., Barrera, R., Boshell, J., & Navarro, J. C. (2004). Venezuelan Equine
1078 Encephalitis. *Annual Review of Entomology*, 49, 141–174.
1079 <https://doi.org/10.1146/annurev.ento.49.061802.123422>

1080 Webster, J. P., Gower, C. M., Knowles, S. C. L., Molyneux, D. H., & Fenton, A. (2016). One health -
1081 an ecological and evolutionary framework for tackling Neglected Zoonotic Diseases.
1082 *Evolutionary Applications*, 9(2), 313–333. <https://doi.org/10.1111/eva.12341>

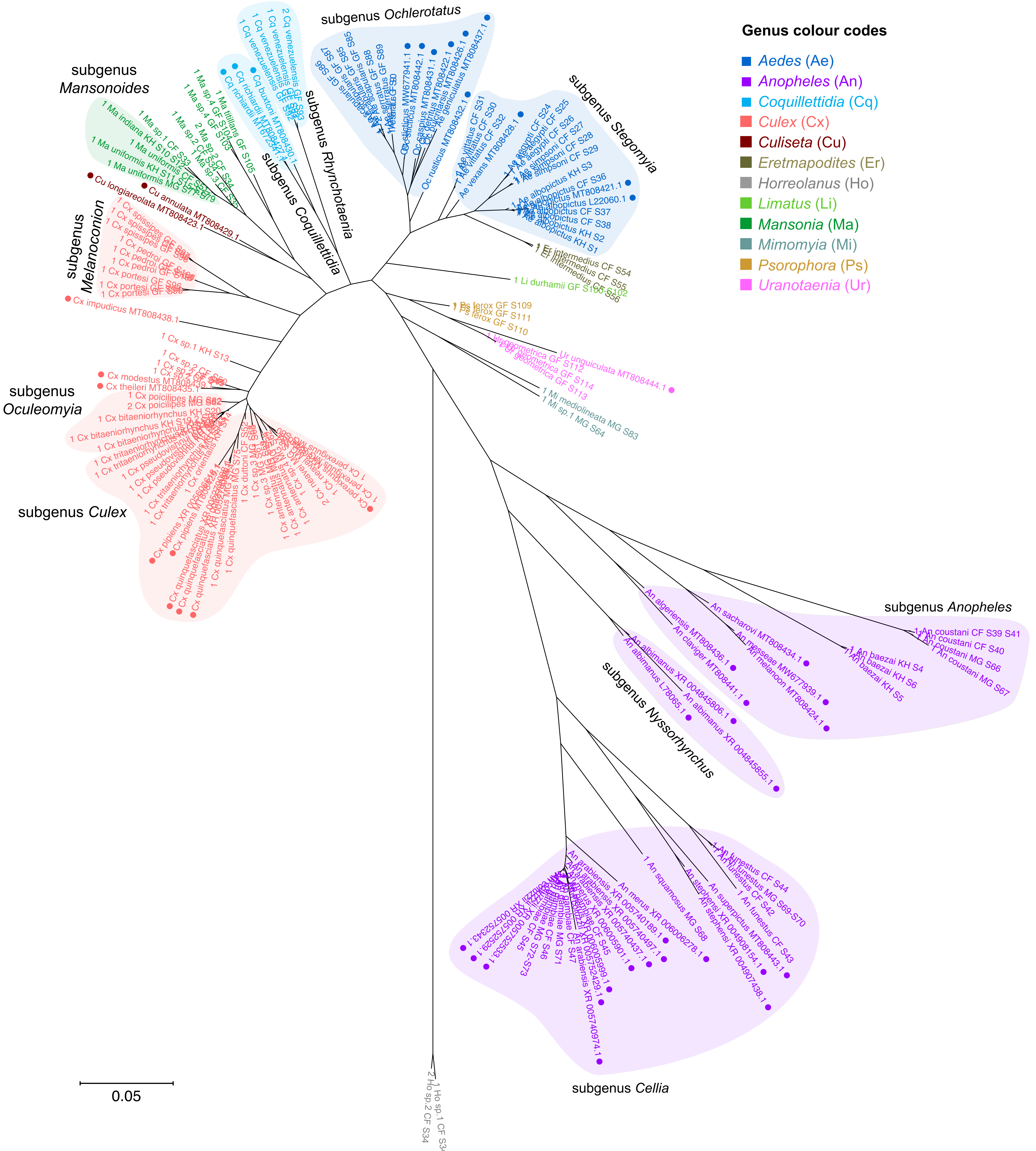
1083 Weedall, G. D., Irving, H., Hughes, M. A., & Wondji, C. S. (2015). Molecular tools for studying the
1084 major malaria vector *Anopheles funestus*: Improving the utility of the genome using a
1085 comparative poly(A) and Ribo-Zero RNAseq analysis. *BMC Genomics*, 16(1), 931.
1086 <https://doi.org/10.1186/s12864-015-2114-z>

1087 WHO. (2017). Global vector control response 2017–2030. In *World Health Organization*.

1088 Zakrzewski, M., Rašić, G., Darbro, J., Krause, L., Poo, Y. S., Filipović, I., Parry, R., Asgari, S., Devine,
1089 G., & Suhrbier, A. (2018). Mapping the virome in wild-caught *Aedes aegypti* from Cairns and
1090 Bangkok. *Scientific Reports*, 8(1), 4690. <https://doi.org/10.1038/s41598-018-22945-y>
1091 Zeller, H., Van Bortel, W., & Sudre, B. (2016). Chikungunya: Its History in Africa and Asia and Its
1092 Spread to New Regions in 2013–2014. *The Journal of Infectious Diseases*, 214(suppl_5), S436–
1093 S440. <https://doi.org/10.1093/INFDIS/JIW391>
1094 Zittra, C., Flechl, E., Kothmayer, M., Vitecek, S., Rossiter, H., Zechmeister, T., & Fuehrer, H. P.
1095 (2016). Ecological characterization and molecular differentiation of *Culex pipiens* complex taxa
1096 and *Culex torrentium* in eastern Austria. *Parasites and Vectors*, 9, 197.
1097 <https://doi.org/10.1186/s13071-016-1495-4>
1098

Genus colour codes

- **Aedes (Ae)**
- **Anopheles (An)**
- **Coquillettidia (Cq)**
- **Culex (Cx)**
- **Culiseta (Cu)**
- **Eretmapodites (Er)**
- **Horreolanus (Ho)**
- **Limatus (Li)**
- **Mansonia (Ma)**
- **Mimomyia (Mi)**
- **Psorophora (Ps)**
- **Uranotaenia (Ur)**



0.05

Ho sp.2 CF S34
Ho sp.1 CF S34