RESEARCH ARTICLE

Novel Wolbachia strains in Anopheles malaria vectors from Sub-Saharan Africa [version 1; referees: 1 approved with reservations]

Claire L. Jeffries1, Gena G. Lawrence2, George Golovko3, Mojca Kristan1, James Orsborne1, Kirstin Spence1, Eliot Hurn1, Janvier Bandibabone1, Luciano M. Tantely1, Fara N. Raharimalala1, Kalil Keita6, Denka Camara6, Yaya Barry6, Francis Wat’senga7, Emile Z. Manzambi7, Yaw A. Afrane8, Abdul R. Mohammed8, Tarekegn A. Abeku9, Shivanand Hedge10, Kamil Khanipov3, Maria Pimenova3, Yuriy Fofanov3, Sebastien Boyer5, Seth R. Irish11, Grant L. Hughes12, Thomas Walker1

1Department of Disease Control, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK
2Entomology Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, Georgia, 30033, USA
3Department of Pharmacology and Toxicology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA
4Laboratoire d’entomologie médicale et parasitologie, Centre de Recherche en Sciences Naturelles (CRSN/LWIRO), Sud-Kivu, Congo, Democratic Republic
5Unité d’Entomologie Médicale, Institut Pasteur de Madagascar, Antananarivo, Madagascar
6Nationale de Lutte contre le Paludisme, Ministere de la Sante, Conakry, Guinea
7National Institute of Biomedical Research, Kinshasa, Congo, Democratic Republic
8Department of Medical Microbiology, University of Ghana, Accra, Ghana
9Malaria Consortium, London, EC2A 4LT, UK
10Department of Pathology, University of Texas Medical Branch, Galveston, Texas, USA
11The US President’s Malaria Initiative and Entomology Branch, Centers for Disease Control and Prevention, Atlanta, Georgia, 30329-4027, USA
12Department of Pathology, Institute for Human Infections and Immunity, Center for Tropical Diseases, Center for Biodefense and Emerging Infectious Disease, University of Texas Medical Branch, Galveston, Texas, USA

Abstract

Background: Wolbachia, a common insect endosymbiotic bacterium that can influence pathogen transmission and manipulate host reproduction, has historically been considered absent from the Anopheles (An.) genera, but has recently been found in An. gambiae s.l. populations. As there are numerous Anopheles species that have the capacity to transmit malaria, we analysed a range of species to determine Wolbachia prevalence rates, characterise novel Wolbachia strains and determine any correlation between the presence of Plasmodium, Wolbachia and the competing endosymbiotic bacterium Asaia.

Methods: Anopheles adult mosquitoes were collected from five malaria-endemic countries: Guinea, Democratic Republic of the Congo (DRC), Ghana, Uganda and Madagascar, between 2013 and 2017. Molecular analysis
of samples was undertaken using quantitative PCR, Sanger sequencing, 
*Wolbachia* multilocus sequence typing (MLST) and high-throughput amplicon 
sequencing of the bacterial 16S rRNA gene.

**Results:** Novel *Wolbachia* strains were discovered in five species: *An. coluzzii*, 
*An. gambiae* s.s., *An. arabiensis*, *An. moucheti* and *An. species 'A'*, increasing 
the number of *Anopheles* species known to be naturally infected. Variable 
prevalence rates in different locations were observed and novel strains were 
phylogenetically diverse, clustering with *Wolbachia* supergroup B strains. We 
also provide evidence for resident strain variants within *An. species 'A'*. 
*Wolbachia* is the dominant member of the microbiome in *An. moucheti* and *An. 
species 'A'*, but present at lower densities in *An. coluzzii*. Interestingly, no 
evidence of *Wolbachia/Asaia* co-infections was seen and *Asaia* infection 
densities were also shown to be variable and location dependent.

**Conclusions:** The important discovery of novel *Wolbachia* strains in 
*Anopheles* provides greater insight into the prevalence of resident *Wolbachia* 
strains in diverse malaria vectors. Novel *Wolbachia* strains (particularly 
high-density strains) are ideal candidate strains for transinfection to create 
stable infections in other *Anopheles* mosquito species, which could be used for 
population replacement or suppression control strategies.

**Keywords**
Wolbachia, mosquitoes, malaria, Anopheles, Asaia, endosymbionts
Background

Malaria is a mosquito-borne disease caused by infection with Plasmodium (P.) parasites, with transmission to humans occurring through the inoculation of Plasmodium sporozoites during blood-feeding of an infectious female Anopheles (An.) mosquito. The genus Anopheles consists of 475 formally recognised species with ~40 vector species/species complexes responsible for the transmission of malaria at a level of public health concern. During the mosquito infection cycle, Plasmodium parasites encounter a variety of resident microbiota both in the mosquito midgut and other tissues. Numerous studies have shown that certain species of bacteria can inhibit Plasmodium development\(^1\). For example, Enterobacter bacteria that reside in the Anopheles midgut can inhibit the development of Plasmodium parasites prior to their invasion of the midgut epithelium\(^9\). Wolbachia endosymbiotic bacteria are estimated to naturally infect ~40% of insect species\(^1\) including mosquito vector species that are responsible for transmission of human diseases, such as Culex (Cx.) quinquefasciatus\(^6\)-\(^10\) and Aedes (Ae.) albopictus\(^11\)-\(^12\). Although Wolbachia strains have been shown to have variable effects on arboviral infections in their native mosquito hosts\(^1\)-\(^5\), transinfected Wolbachia strains have been considered for mosquito biocontrol strategies, due to observed arbovirus transmission blocking abilities and a variety of synergistic phenotypic effects. Transinfected strains in Ae. aegypti and Ae. albopictus provide strong inhibitory effects on arboviruses, with maternal transmission and cytoplasmic incompatibility enabling introduced strains to spread through populations\(^6\)-\(^12\). Open releases of Wolbachia-transinfected Ae. aegypti populations have demonstrated the ability of the wMel Wolbachia strain to invade wild populations\(^13\) and provide strong inhibitory effects on viruses from field populations\(^14\), with releases currently occurring in arbovirus endemic countries such as Indonesia, Vietnam, Brazil and Colombia (https://www.worldmosquitoprogram.org).

The prevalence of Wolbachia in Anopheles species has not been extensively studied, with most studies focused in Asia using classical PCR-based screening; up until 2014 there was no evidence of resident strains in mosquitoes from this genus\(^13\)-\(^20\). Furthermore, significant efforts to establish artificially infected lines were, up until recently, also unsuccessful\(^20\). Somatic, transient infections of the Wolbachia strains wMelPop and wAlbB in An. gambiae were shown to significantly inhibit P. falciparum\(^1\), but the interference phenotype is variable with other Wolbachia strain-parasite combinations\(^2\)-\(^4\). A stable line was established in An. stephensi, a vector of malaria in southern Asia, using the wAlbB strain and this was also shown to confer resistance to P. falciparum infection\(^13\). One potential reason postulated for the absence of Wolbachia in Anopheles species was thought to be the presence of other endosymbiotic bacteria, particularly from the genus Asaia\(^9\). This acetic acid bacterium is stably associated with several Anopheles species and is often the dominant species in the mosquito microbiota\(^7\). In laboratory studies, Asaia has been shown to impede the vertical transmission of Wolbachia in Anopheles\(^8\) and was shown to have a negative correlation with Wolbachia in mosquito reproductive tissues\(^8\).

Recently, resident Wolbachia strains have been discovered in the An. gambiae s.l. complex, which consists of multiple morphologically indistinguishable species including several major malaria vector species. Wolbachia strains (collectively named wAnga) were found in An. gambiae s.l. populations in Burkina Faso\(^9\) and Mali\(^8\), suggesting that Wolbachia may be more abundant in the An. gambiae complex across Sub-Saharan Africa. Globally, there is a large variety of Anopheles vector species (~70) that have the capacity to transmit malaria\(^1\) and could potentially contain resident Wolbachia strains. Additionally, this number of malaria vector species may be an underestimate given that recent studies using molecular barcoding have also revealed a larger diversity of Anopheles species than would have be identified using morphological identification alone\(^2\)-\(^4\).

In this study, we collected Anopheles mosquitoes from five malaria-endemic countries; Ghana, Democratic Republic of the Congo (DRC), Guinea, Uganda and Madagascar, from 2013–2017. Wild-caught adult female Anopheles were screened for P. falciparum malaria parasites, Wolbachia and Asaia bacteria. In total, we analysed mosquitoes from 17 Anopheles species that are known malaria vectors or implicated in transmission, and some unidentified species, discovering five species of Anopheles with resident Wolbachia strains; An. coluzzii from Ghana, An. gambiae s.s., An. arabiensis, An. moucheti and An. species ‘A’ from DRC. Using Wolbachia gene sequencing, including multilocus sequence typing (MLST), we show that the resident strains in these malaria vectors are diverse, novel strains and quantitative PCR (qPCR) and 16S rRNA amplicon sequencing data suggests that the strains in An. moucheti and An. species ‘A’ are higher density infections, compared to the strains found in the An. gambiae s.l. complex. We found no evidence for either Wolbachia-Asaia co-infections, or for either endosymbiont having any significant effect on the prevalence of Plasmodium in wild mosquito populations.

Methods

Study sites & collection methods

Anopheles adult mosquitoes were collected from five malaria-endemic countries in Sub-Saharan Africa (Guinea, Democratic Republic of the Congo (DRC), Ghana, Uganda and Madagascar) between 2013 and 2017 (Figure 1). Human landing catches, CDC light traps and pyrethrum spray catch were undertaken between April 2014 and February 2015 in 10 villages near four cities in Guinea; Foulahay (10.144633, -10.749717) and Balayani (10.1325, -10.7443) near Faranah; Djoumay (10.836317, -14.2481) and Kaboye Amara (10.93435, -14.36995) near Boke; Tongbekoro (9.294295, -10.147953), Keredou (9.208919, -10.069525), and Gbhangbadou (9.274363, -9.998639) near Kissidougou; and Makonon (10.291124, -9.363358), Balandou (10.407669, -9.219096), and Dalabani (10.463692, -9.451904) near Kankan. Human landing catches and pyrethrum spray catches were undertaken between January and September 2015 in seven sites of the DRC; Kinshasa (-4.415881, 15.412188), Mikalayi (-6.024184, 22.318251), Kisangmani (0.516350, 25.221176), Katana (-2.225129, 28.831604), Kalemie (-5.919054, 29.186572), and Kapolowe (-10.939802, 26.952970). We also analysed a subset from collections obtained from Lwiro (-2.244097, 28.815232),

Page 3 of 22

Wellcome Open Research 2018, 3:113 Last updated: 23 SEP 2018
Figure 1. Locations of Anopheles species collections (including Wolbachia-infected species) and P. falciparum malaria prevalence rates in mosquitoes (across all species for each location). (A) Overall map showing the five malaria-endemic countries where mosquito collections were undertaken. (B) High P. falciparum prevalence rates in Guinea, and Wolbachia-infected An. coluzzii from Ghana (no P. falciparum detected). (C) Wolbachia strains in An. gambiae s.s., An. arabiensis, An. species A and An. moucheti from Democratic Republic of the Congo (DRC) and variable P. falciparum prevalence rate in DRC and Uganda. (D) Low P. falciparum infection rates in Madagascar and no evidence of resident Wolbachia strains. (W+: Wolbachia detected in this species).

a village near Katana, collected between September and October 2015. A combination of CDC light traps, pyrethrum spray catches and human landing catches were undertaken in Butemba, Kyankwanzi District in mid-western Uganda (1.1068444, 31.5910085) in August and September 2013, and June 2014. CDC light trap catches were undertaken in May 2017 in Dogo in Ada, Greater Accra, Ghana (5.874861111, 0.560611111). In Madagascar, sampling was undertaken in June 2016 at four sites: Anivorano Nord, located in the Northern domain, (-12.7645000, 49.2386944); Ambomiharina, Western domain, (-16.3672778, 46.9928889); Antafia, Western domain, (-17.0271667, 46.7671389); and Ambohimarina, Central domain, (-18.3329444, 47.1092500). Trapping consisted of CDC light traps and a net trap baited with Zebu (local species of cattle) to attract zoophilic species.

DNA extraction and species identification
DNA was extracted from individual whole mosquitoes or abdomens using QIAGEN DNeasy Blood and Tissue Kits according to manufacturer’s instructions. DNA extracts were eluted in a final volume of 100 μl and stored at −20°C. Species identification was initially undertaken using morphological keys followed by diagnostic species-specific PCR assays to distinguish between the morphologically indistinguishable sibling mosquito species of the An. gambiae complex and An. funestus complex.

To determine species identification for samples of interest and samples that could not be identified by species-specific PCR, Sanger sequences were generated from ITS2 PCR products.

Detection of P. falciparum and Asaia
Detection of P. falciparum malaria was undertaken using qPCR targeting an 120-bp sequence of the P. falciparum cytochrome c oxidase subunit 1 (Cox1) mitochondrial gene as preliminary trials revealed this was the optimal method for both sensitivity and specificity. Positive controls from gDNA extracted from a cultured P. falciparum-infected blood sample (parasitaemia of ~10%) were serially diluted to determine the threshold
limit of detection, in addition to the inclusion no template controls (NTCs). *Asaia* detection was undertaken targeting the 16S rRNA gene\(^{31}\). Ct values for both *P. falciparum* and *Asaia* assays in selected *An. gambiae* extracts were normalized to Ct values for a single copy *An. gambiae* rps17 housekeeping gene (accession no. AGAP004887 on www.vectorbase.org)\(^{23,35}\). As Ct values are inversely related to the amount of amplified DNA, a higher target gene Ct: host gene Ct ratio represented a lower estimated infection level. qPCR reactions were prepared using 5 μl of FastStart SYBR Green Master mix (Roche Diagnostics), a final concentration of 1 μM of each primer, 1 μl of PCR grade water and 2 μl template DNA, to a final reaction volume of 10 μl. Prepared reactions were run on a Roche LightCycler® 96 System and amplification was followed by a dissociation curve (95°C for 10 seconds, 65°C for 60 seconds and 97°C for 1 second) to ensure the correct target sequence was being amplified. PCR results were analysed using the LightCycler® 96 software (Roche Diagnostics). A sub-selection of PCR products from each assay was sequenced to confirm correct amplification of the target gene fragment.

**Wolbachia detection**

*Wolbachia* detection was first undertaken targeting three conserved *Wolbachia* genes previously shown to amplify a wide diversity of strains; 16S rDNA gene\(^{31,34}\), *Wolbachia* surface protein (*wsp*) gene\(^{33}\) and *FtsZ* cell cycle gene\(^{35}\). DNA extracted from a *Drosophila melanogaster* fly (infected with the wMel strain of *Wolbachia*) was used a positive control, in addition to no template controls (NTCs). The 16S rDNA\(^{31}\) and *wsp*\(^{33}\) gene PCR reactions were carried out in a Bio-Rad T100 Thermal Cycler using standard cycling conditions and PCR products were separated and visualised using 2% E-Gel EX agarose gels (Invitrogen) with SYBR safe and an Invitrogen E-Gel iBase Real-Time Transilluminator. *FtsZ*\(^{35}\) and 16S rDNA\(^{31}\) gene real time PCR reactions were prepared using 5 μl of FastStart SYBR Green Master mix (Roche Diagnostics), a final concentration of 1 μM of each primer, 1 μl of PCR grade water and 2 μl template DNA, to a final reaction volume of 10 μl. Prepared reactions were run on a Roche LightCycler® 96 System for 15 minutes at 95°C, followed by 40 cycles of 95°C for 15 seconds and 58°C for 30 seconds. Amplification was followed by a dissociation curve (95°C for 10 seconds, 65°C for 60 seconds and 97°C for 1 second) to ensure the correct target sequence was being amplified. PCR results were analysed using the LightCycler® 96 software (Roche Diagnostics). To estimate *Wolbachia* densities across multiple *Anopheles* mosquito species, *ftsZ* and 16S qPCR Ct values were compared to total dsDNA extracted, measured using an Invitrogen Qubit 4 fluorometer. A serial dilution series of a known *Wolbachia* concentration was being amplified. PCR products were separated and visualised using 2% E-Gel EX agarose gels (Invitrogen) with SYBR safe and an Invitrogen E-Gel iBase Real-Time Transilluminator. PCR products were submitted to Source BioScience (Source BioScience Plc, Nottingham, UK) for PCR reaction clean-up, followed by Sanger sequencing to generate both forward and reverse reads. Sequencing analysis was carried out in MEGA\(^{78}\) as follows. Both chromatograms (forward and reverse traces) from each sample was manually checked, edited, and trimmed as required, followed by alignment with ClustalW and checking to produce consensus sequences. Consensus sequences were used to perform nucleotide BLAST (NCBI) database queries, and searches against the *Wolbachia* MLST database\(^{36}\). If a sequence produced an exact match in the MLST database we assigned the appropriate allele number, otherwise we obtained a new allele number for each novel gene locus sequence through submission of the FASTA and raw trace files on the *Wolbachia* MLST website for new allele assignment and inclusion within the database. Full consensus sequences were also submitted to GenBank and assigned accession numbers. The Sanger sequencing traces from the *wsp* gene were also treated in the same way and analysed alongside the MLST gene locus scheme, as an additional marker for strain typing.

**Phylogenetic analysis**

Alignments were constructed in MEGA7 by ClustalW to include all relevant and available sequences highlighted through searches on the BLAST and *Wolbachia* MLST databases. Maximum Likelihood phylogenetic trees were constructed from Sanger sequences as follows. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model\(^{40}\). The tree with the highest log likelihood in each case is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automati-

**Microbiome analysis**

The microbiomes of selected individual *Anopheles* were analysed using barcoded high-throughput amplicon sequencing of the bacterial 16S rRNA gene. Sequencing libraries for each isolate were generated using universal 16S rRNA V3-V4 region primers\(^{61}\) in accordance with Illumina 16S rRNA previously optimised protocols\(^{37}\). A *Cx. pipiens* gDNA extraction (previously shown to be infected with the wPip strain of *Wolbachia*) was used a positive control for each PCR run, in addition to no template controls (NTCs). If no amplification was detected using standard primers, further PCR analysis was undertaken using degenerate primers\(^{37}\). PCR products were separated and visualised using 2% E-Gel EX agarose gels (Invitrogen) with SYBR safe and an Invitrogen E-Gel iBase Real-Time Transilluminator. PCR products were submitted to Source BioScience (Source BioScience Plc, Nottingham, UK) for PCR reaction clean-up, followed by Sanger sequencing to generate both forward and reverse reads. Sequencing analysis was carried out in MEGA\(^{78}\) as follows. Both chromatograms (forward and reverse traces) from each sample was manually checked, edited, and trimmed as required, followed by alignment with ClustalW and checking to produce consensus sequences. Consensus sequences were used to perform nucleotide BLAST (NCBI) database queries, and searches against the *Wolbachia* MLST database\(^{36}\). If a sequence produced an exact match in the MLST database we assigned the appropriate allele number, otherwise we obtained a new allele number for each novel gene locus sequence through submission of the FASTA and raw trace files on the *Wolbachia* MLST website for new allele assignment and inclusion within the database. Full consensus sequences were also submitted to GenBank and assigned accession numbers. The Sanger sequencing traces from the *wsp* gene were also treated in the same way and analysed alongside the MLST gene locus scheme, as an additional marker for strain typing.

**Microbiome analysis**

The microbiomes of selected individual *Anopheles* were analysed using barcoded high-throughput amplicon sequencing of the bacterial 16S rRNA gene. Sequencing libraries for each isolate were generated using universal 16S rRNA V3-V4 region primers\(^{61}\) in accordance with Illumina 16S rRNA
metagenomic sequencing library protocols. The samples were barcoded for multiplexing using Nextera XT Index Kit v2. Sequencing was performed on an Illumina MiSeq instrument using a MiSeq Reagent Kit v2 (500-cycles). Quality control and taxonomical assignment of the resultant reads were performed using CLC Genomics Workbench 8.0.1 Microbial Genomics Module. Low quality reads containing nucleotides with quality threshold below 0.05 (using the modified Richard Mott algorithm), as well as reads with two or more unknown nucleotides were removed from analysis. Additionally, reads were trimmed to remove sequenced Nextera adapters. Reference-based operational taxonomic unit (OTU) picking was performed using the SILVA SSU v128 97% database. Sequences present in more than one copy but not clustered to the database were then placed into de novo OTUs (97% similarity) and aligned against the reference database with 80% similarity threshold to assign the “closest” taxonomical name where possible. Chimeras were removed from the dataset if the absolute crossover cost was 3 using a k-mer size of 6. Alpha diversity was measured using Shannon entropy (OTU level).

Statistical analysis
Fisher’s exact post hoc test in Graphpad Prism 7 was used to compare infection rates. Normalised qPCR Ct ratios were compared using unpaired t-tests in GraphPad Prism 7.

Results
Mosquito species and resident Wolbachia strains
Anopheles species composition varied depending on country and mosquito collection sites (Table 1). We detected Wolbachia

<table>
<thead>
<tr>
<th>Country</th>
<th>Location</th>
<th>Species</th>
<th>Individuals</th>
<th>Infection prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P. falciparum</td>
</tr>
<tr>
<td>Guinea</td>
<td>Faranah</td>
<td>An. gambia s.s.</td>
<td>48</td>
<td>43.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. nili</td>
<td>9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Kissidougou</td>
<td>An. gambia s.s.</td>
<td>44</td>
<td>18.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. species O</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Boke</td>
<td>An. gambia s.s.</td>
<td>21</td>
<td>52.4</td>
</tr>
<tr>
<td></td>
<td>Kankan</td>
<td>An. gambia s.s.</td>
<td>48</td>
<td>38.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. sp. unknown</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>DRC</td>
<td>Mikalayi</td>
<td>An. gambia s.s.</td>
<td>16</td>
<td>29.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. moucheti</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. funestus s.s.</td>
<td>13</td>
<td>30.8</td>
</tr>
<tr>
<td></td>
<td>Kisangani</td>
<td>An. gambia s.s.</td>
<td>25</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>4</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>Katana</td>
<td>An. gambia s.s.</td>
<td>23</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. funestus s.s.</td>
<td>5</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. species A</td>
<td>2</td>
<td>50.0</td>
</tr>
<tr>
<td></td>
<td>Lwiro (Katana)</td>
<td>An. species A</td>
<td>33</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Kapolowe</td>
<td>An. gambia s.s.</td>
<td>9</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. funestus s.s.</td>
<td>5</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>Kalemie</td>
<td>An. gambia s.s.</td>
<td>28</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Kinshasa</td>
<td>An. gambia s.s.</td>
<td>27</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. funestus s.s.</td>
<td>2</td>
<td>50.0</td>
</tr>
<tr>
<td>Ghana</td>
<td>Dogo</td>
<td>An. coluzzii</td>
<td>286</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. melas</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Uganda</td>
<td>Butemba (2013)</td>
<td>An. gambia s.s.</td>
<td>57</td>
<td>19.3</td>
</tr>
<tr>
<td></td>
<td>Butemba (2014)</td>
<td>An. gambia s.s.</td>
<td>135</td>
<td>36.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Country</td>
<td>Location</td>
<td>Species</td>
<td>Individuals</td>
<td>Infection prevalence (%)</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>------------------</td>
<td>-------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P. falciparum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. funestus</td>
<td>8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. gambiae s.s.</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>Anivorano</td>
<td>Nord</td>
<td>An. maculipalpis</td>
<td>38</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. coustani</td>
<td>9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. rufipes</td>
<td>22</td>
<td>0.0</td>
</tr>
<tr>
<td>Madagascar</td>
<td></td>
<td>An. gambiae s.s.</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. maculipalpis</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. coustani</td>
<td>11</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. rufipes</td>
<td>11</td>
<td>0.0</td>
</tr>
<tr>
<td>Ambomiharina</td>
<td></td>
<td>An. funestus</td>
<td>11</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. pharoensis</td>
<td>7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. rufipes</td>
<td>21</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. maculipalpis</td>
<td>9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. gambiae s.s.</td>
<td>8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. coustani</td>
<td>24</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. squamosus</td>
<td>10</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. mascarensis</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. pauliani</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td>Antafia</td>
<td></td>
<td>An. gambiae s.s.</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. pauliani</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. rufipes</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. mascarensis</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>Ambohimarina</td>
<td></td>
<td>An. funestus</td>
<td>11</td>
<td>27.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. gambiae s.s.</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. arabiensis</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. rufipes</td>
<td>7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. coustani</td>
<td>18</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. maculipalpis</td>
<td>8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. squamosus</td>
<td>46</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An. mascarensis</td>
<td>11</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Adult individuals from Lwiro (Katana), DRC were collected as both larvae and adults so have been excluded from P. falciparum and Asaia prevalence analysis.

in *An. coluzzii* (previously named M molecular form) mosquitoes from Ghana (prevalence of 4% - termed wAnga-Ghana) and *An. gambiae* s.s. (previously named S molecular form) from all six collection sites in DRC (prevalence range of 8–24%) in addition to a single infected *An. arabiensis* from Kalemie in DRC (Figure 1 and Table 1). The molecular phylogeny of the ITS2 gene of *Anopheles gambiae* s.l. complex individuals (including both *Wolbachia*-infected and uninfected individuals analysed in our study) confirmed molecular species identifications made using species-specific PCR assays (Figure 2). Novel resident *Wolbachia* infections were detected in two additional *Anopheles* species from DRC; *An. moucheti* (termed wAnM) from Mikalayi, and *An. species A* (termed wAnsA) from Katana. Additionally, we screened adult female mosquitoes of *An. species A* (collected as larvae and adults) from Lwiro, a village near Katana in DRC, and detected *Wolbachia* in 30/33 (91%), indicating this resident wAnsA strain has a high infection prevalence in populations in this region. The molecular phylogeny of the ITS2 gene revealed *Wolbachia*-infected individuals from Lwiro and Katana are the same *An. species A* (Figure 3) previously collected in Eastern Zambia43 and Western Kenya44. All ITS2 sequences were deposited in GenBank (accession numbers MH598414–MH598445; listed in Supplementary Table 1).

**Wolbachia** strain typing

Phylogenetic analysis of the 16S rRNA gene demonstrated that the 16S sequences for these strains cluster with other Super-group B strains such as wPip (99–100% nucleotide identity) (Figure 4a). When compared to the resident *Wolbachia* strains
Figure 2. Maximum Likelihood molecular phylogenetic analysis of Anopheles gambiae complex ITS2 sequences from field-collected mosquitoes. The tree with the highest log likelihood (-785.65) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 42 nucleotide sequences. There were a total of 475 positions in the final dataset. W+; individual was Wolbachia positive, W-; individual was Wolbachia negative. DRC, Democratic Republic of the Congo (red); KAL, Kalemie; MIK, Mikalayi; KIN, Kinshasa; KAT, Katana; GHA, Ghana (blue); DOG, Dogo; GUI, Guinea (green); KSK, Kissidougou; MAD, Madagascar (purple); ANT, Antafia; UGA, Uganda (maroon); BUT, Butemba.
Figure 3. Maximum Likelihood molecular phylogenetic analysis of Anopheles ITS2 sequences from field-collected mosquitoes outside of the An. gambiae s.l. complex. The tree with the highest log likelihood (-3084.12) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 118 nucleotide sequences. There were a total of 156 positions in the final dataset. W+; individual was Wolbachia positive, W-; individual was Wolbachia negative. DRC, Democratic Republic of the Congo (red); KAT, Katana; LWI, Lwiro; MIK, Mikalayi; GUI, Guinea (green); FAR, Faranah; KAN, Kankan; KSK, Kissidougou; MAD, Madagascar (purple); AMB, Ambomiharina.
Figure 4. Resident Wolbachia strain phylogenetic analysis using 16S rRNA and wsp genes. (A) Maximum Likelihood molecular phylogenetic analysis of the 16S rRNA gene for resident strains in An. coluzzii (wAnga-Ghana; blue), An. moucheti (wAnM; green) and An. species A (wAnsA; red). The tree with the highest log likelihood (-660.03) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 17 nucleotide sequences. There were a total of 333 positions in the final dataset. Accession numbers of additional sequences obtained from GenBank are shown, including wPip (navy blue), wAnga-Mali (purple) and wAnga-Burkina Faso strains (maroon). (B) Maximum Likelihood molecular phylogenetic analysis of the wsp gene for wAnsA-infected representative individuals from the DRC (red). The tree with the highest log likelihood (-3663.41) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 83 nucleotide sequences. There were a total of 443 positions in the final dataset. Reference numbers of additional sequences obtained from the MLST database or GenBank (accession number) are shown. Strains isolated from mosquitoes are highlighted in navy blue. KAT, Katana; LWI, Lwiro; IsoN, Isolate number.
in *An. gambiae* s.l. populations from Mali\(^{60}\) and Burkina Faso\(^{68}\), *wAnga-Ghana* is more closely related to the Supergroup B strain of *wAnga* from Burkina Faso. Although a resident strain was detected in *An. gambiae* s.s. and a single *An. arabiensis* from DRC through amplification of 16S rRNA fragments using two independent PCR assays\(^{64,65}\), we were unable to obtain 16S sequences of sufficient quality to allow further analysis. The *Wolbachia* wsp gene has been evolving at a faster rate and provides more informative strain phylogenies\(^{45}\). As expected, however, and similar to *Wolbachia*-infected *An. gambiae* s.l. from Burkina Faso\(^{57}\) and Mali\(^{70}\), a fragment of the wsp gene was not amplified from *Wolbachia*-positive samples from *An. gambiae* s.s., *An. arabiensis* and *An. coluzzii*. Similarly, no wsp gene fragment amplification occurred from *wAnM*-infected *An. mouchetti*. However, wsp sequences were obtained from both *Wolbachia*-infected individuals of *An. species A* from Katana. We also analysed the wsp sequences of 22 specimens of *An. species A* from Lwiro (near Katana) and found identical sequences to the two individuals from Katana. Phylogenetic analysis of the wsp sequences obtained for the *wAnS* strain, for both individuals from Katana (*wAnS* wsp DRC-KAT1, *wAnS* wsp DRC-KAT2) and three representative individuals from Lwiro (*wAnS* wsp DRC-LWI1, *wAnS* wsp DRC-LWI2, *wAnS* wsp DRC-LWI3) indicates *wAnS* is most closely related to *Wolbachia* strains of Supergroup B (such as *wPip*, *wAlbB*, *wMa* and *wNo*), which is consistent with 16S rRNA phylogeny. However, the improved phylogenetic resolution provided by wsp indicates they cluster separately (Figure 4b). Typing of the *wAnS* wsp nucleotide sequences highlighted that there were no exact matches to wsp alleles currently in the *Wolbachia* MLST database. All *Wolbachia* 16S and wsp sequences of sufficient quality to generate a consensus were deposited into GenBank (accession numbers MH605275–MH605285; listed in Supplementary Table 2).

MLST was undertaken to provide more accurate strain phylogenies. This was done for the novel *Wolbachia* strains *wAnM* and *wAnS* in addition to the resident *wAnga-Ghana* strain in *An. coluzzii* from Ghana. We were unable to amplify any of the five MLST genes from *Wolbachia*-infected *An. gambiae* s.s. and *An. arabiensis* from DRC (likely due to low infection densities). New alleles for all five MLST gene loci (sequences differed from those currently present in the MLST database) and novel allelic profiles confirm the diversity of these novel *Wolbachia* strains (Table 2). The phylogeny of these three novel strains based on concatenated sequences of all five MLST gene loci confirms they cluster within Supergroup B (Figure 5a). This also demonstrates the novelty as comparison with a wide range of strains (including all isolates highlighted through partial matching during typing of each locus) shows these strains are distinct from currently available sequences (Figure 5a and Table 2). The concatenated phylogeny indicates that *wAnM* is most closely related to a Hemiptera strain: Isolate number 1616 found in *Bemisia tabaci* in Uganda, and a Coleoptera strain: Isolate number 20 found in *Tribolium confusum*. Concatenation of the MLST loci also indicates *wAnS* is closest to a group containing various Lepidoptera and Hymenoptera strains from multiple countries in Asia, Europe and America, as well as two mosquito strains: Isolate numbers 1830 and 1831, found in *Aedes cinereus* and *Coquillettidia richiardii* in Russia. This highlights the lack of concordance between *Wolbachia* strain phylogeny and their insect hosts across diverse geographical regions. We also found evidence of potential strain variants in *wAnS* through variable MLST gene fragment amplification and resulting closest-match allele numbers.

A second *wAnS*-infected sample, *An. sp. A/1 (W+)* DRC-KAT2, only successfully amplified *hcpA* and *coxA* gene fragments and although identical sequences were obtained for wsp (Figure 4b) and *hcpA*, genetic diversity was seen in the *coxA* sequences, with typing indicating a different, but still novel allele for the *coxA* sequence from this individual (*wAnS*(2) *coxA* DRC-KAT2) (Figure 5b). Further analysis of the *coxA* sequence as part of MLST allele submission from this variant suggested the possibility of a double infection, where two differing strains of *Wolbachia* are present. MLST gene fragment amplification was also variable for *wAnga-Ghana*-infected *An. coluzzii*, requiring two individuals to generate the five

<table>
<thead>
<tr>
<th>Mosquito species</th>
<th>Wolbachia strain</th>
<th><em>Wolbachia</em> MLST gene allele numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>gatB</em></td>
</tr>
<tr>
<td>An. species A</td>
<td><em>wAnS</em></td>
<td>279</td>
</tr>
<tr>
<td>An. mouchetti</td>
<td><em>wAnM</em></td>
<td>280</td>
</tr>
<tr>
<td>An. coluzzii</td>
<td><em>wAnga-Ghana</em></td>
<td>9</td>
</tr>
</tbody>
</table>

*Alternative degenerate primers (set 3) were used to generate sequence from another *An. coluzzii* individual from the same location to complete the full allelic profile.

Table 2. Novel resident *Wolbachia* strain multilocus sequence typing (MLST) gene allelic profiles. Novel allele numbers (in bold) assigned by the *Wolbachia* MLST database for strains from *An. species A* (*wAnS*) and *An. mouchetti* (*wAnM*) are shown, alongside the novel allelic profile from *An. coluzzii* (*wAnga-Ghana*), comprising exact matches to existing alleles present in the database for each gene locus.
Figure 5. *Wolbachia* multilocus sequence typing (MLST) phylogenetic analysis of resident *Wolbachia* strains in *An. coluzzii*, *An. moucheti* and *An. species A.* (A) Maximum Likelihood molecular phylogenetic analysis from concatenation of all five MLST gene loci for resident *Wolbachia* strains from *An. coluzzii* (*wAnga-Ghana*; blue), *An. moucheti* (*wAnM*; green) and *An. species A* (*wAnsA*; red). The tree with the highest log likelihood (-10606.13) is shown and drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 94 nucleotide sequences. There were a total of 2067 positions in the final dataset. Concatenated sequence data from *Wolbachia* strains downloaded from MLST database for comparison shown with isolate numbers in brackets (IsoN). *Wolbachia* strains isolated from mosquito species highlighted in navy blue, bold. Strains isolated from other Dipteran species are shown in navy blue, from Coleoptera in olive green, from Hemiptera in purple, from Hymenoptera in teal blue, from Lepidoptera in maroon and from other, or unknown orders in black. (B) Maximum Likelihood molecular phylogenetic analysis for coxA gene locus for resident *Wolbachia* strains from *An. coluzzii* (*wAnga-Ghana*; blue), *An. moucheti* (*wAnM*; green) and *An. species A* (*wAnsA* and *wAnsA(2)*; red). The tree with the highest log likelihood (-1921.11) is shown and drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 84 nucleotide sequences. There were a total of 402 positions in the final dataset. Sequence data for the coxA locus from *Wolbachia* strains downloaded from MLST database for comparison shown in black and navy blue with isolate numbers (IsoN) from MLST database shown in brackets. *Wolbachia* strains isolated from mosquito species highlighted in navy blue. GenBank sequence for *wAnga-Mali* coxA shown in maroon with accession number.
MLST gene sequences, and for the hcpA locus, more degenerate primers (hcpA_F3/hcpA_R3) were required to generate sequence of sufficient quality for analysis. This is likely due to the low density of this strain potentially influencing the ability to successfully amplify all MLST genes, in addition to the possibility of genetic variation in primer binding regions. Despite the sequences generated for this strain producing exact matches with alleles in the database for each of the five gene loci, the resultant allelic profile, and therefore strain type, did not produce a match, showing this wAnga-Ghana strain is also a novel strain type. The closest matches to the wAnga-Ghana allelic profile were with strains from two Lepidopteran species: Isolate number 609 found in Fabriciana adippe from Russia, and Isolate number 658 found in Pammene fasciana from Greece, but each of these only produced a match for three out of the five loci. The concatenated phylogeny for this strain (Figure 5a) indicates that across the 5 MLST loci, wAnga-Ghana is actually most closely related to a Lepidopteran strain found in Thersamonia thersamon in Russia (Isolate number 132).

Resident strain densities and relative abundance

The relative densities of Wolbachia strains were estimated using qPCR targeting the ftsZ[46] and 16S rRNA[40] genes. qPCR analysis of ftsZ and 16S rRNA indicated the amount of Wolbachia detected in wAnsA-infected and wAnM-infected females was approximately 1000-fold higher (Ct values 20-22) than Wolbachia-infected An. gambiae s.s., An. arabiensis and wAnga-Ghana-infected An. coluzzii (Ct values 30-33). To account for variation in mosquito body size and DNA extraction efficiency, we compared the total amount of DNA for Wolbachia-infected mosquito extracts and conversely, we found less total DNA in the wAnsA-infected extract (1.36 ng/μl) and the An. moucheti (wAnM-infected) extract (5.85 ng/μl) compared to the mean of 6.64 ± 2.33 ng/μl for wAnga-Ghana-infected An. coluzzii. To estimate the relative abundance of resident Wolbachia strains in comparison to other bacterial species, we sequenced the bacterial microbiome using 16S rRNA amplicon sequencing on Wolbachia-infected individuals. We found wAnsA, wAnsA(2) and wAnM Wolbachia strains were the dominant OTUs of these mosquito species (Figure 6). In contrast, the lower-density infection wAnga-Ghana strain represented only ~10% of the OTUs within the microbiome.

P. falciparum, Wolbachia and Asaia prevalence

The prevalence of P. falciparum in female mosquitoes was extremely variable across countries and collection locations (Figure 1 and Table 1) with very high prevalence recorded in An. gambiae s.s. from villages close to Boke (52%) and Faranah (44%) in Guinea. Despite the collection of other Anopheles species in Guinea, An. gambiae s.s. was the only species to have detectable malaria parasite infections. In contrast, P. falciparum was detected in multiple major vector species from DRC, including An. gambiae s.s., An. arabiensis and An. funestus s.s. A high prevalence of P. falciparum was also detected in An. gambiae s.s. from Uganda for both collection years; 19% for 2013 and 36% for 2014. In contrast, no P. falciparum infections were detected in any of the An. coluzzii or An. melas collected in Ghana. In Madagascar, P. falciparum was detected in only two species; An. gambiae s.s. and An. rufipes. We compared the

![Figure 6. The relative abundance of resident Wolbachia strains in Anopheles.](image-url)
overall *P. falciparum* infection rates in *An. gambiae* s.s. mosquitoes collected across all locations from DRC to determine if there was any correlation with the presence of the low density *w*Anga-DRC *Wolbachia* resident strain. Overall, of the 128 mosquitoes collected, only 1.56% (n=2) had detectable *Wolbachia-Plasmodium* co-infections, compared to 10.16% (n=13) where we only detected *Wolbachia*. A further 11.72% (n=15) were only PCR-positive for *P. falciparum*. As expected, for the vast majority of mosquitoes (76.56%, n=98) we found no evidence of *Wolbachia* or *P. falciparum* present, resulting in no correlation across all samples (Fisher’s exact post hoc test on unnormalized data, two-tailed, P=0.999). Interestingly, one *An. species ‘A’* female from Katana was infected with *P. falciparum*.

For all *Wolbachia*-infected females collected in our study (including *An. coluzzii* from Ghana and novel resident strains in *An. moucheti* and *An. species A*), we did not detect the presence of *Asaia*. No resident *Wolbachia* strain infections were detected in *Anopheles* mosquitoes from Guinea, Uganda or Madagascar. However, high *Asaia* and malaria parasite prevalence rates were present in *Anopheles* mosquitoes from Guinea and Uganda (including multiple species in all four sites in Guinea). We compared the overall *P. falciparum* infection rates in *An. gambiae* s.s. collected across all locations from Guinea, with and without *Asaia* bacteria, and found no overall correlation (Fisher’s exact post hoc test on unnormalized data, two-tailed, P=0.4902). There was also no overall correlation between *Asaia* and *P. falciparum* infections in *An. gambiae* s.s. from Uganda for both 2013 (Fisher’s exact post hoc test on unnormalized data, two-tailed, P=0.601) and 2014 (Fisher’s exact post hoc test on unnormalized data, two-tailed, P=0.282).

*Asaia* can be environmentally acquired at all life stages but can also have the potential to be vertically and horizontally transmitted between individual mosquitoes. Therefore, we performed 16S microbiome analysis on a sub-sample of *Asaia*-infected *An. gambiae* s.s. from Kissidougou (Guinea), a location in which high levels of *Asaia* were detected by qPCR (mean *Asaia* Ct = 17.84 ± 2.27)44. *Asaia* in these individuals is the dominant bacterial species present (Figure 7a) but in Uganda we detected much lower levels of *Asaia* (qPCR mean Ct = 33.33 ± 0.19) and this was reflected in *Asaia* not being a dominant species (Figure 7b). The alpha and beta diversity of *An. gambiae* s.s. from Kissidougou, Guinea and Butembo, Uganda shows much more overall diversity in the microbiome for Uganda individuals (Supplementary Figure 1). Interestingly, 2/5 of these individuals from Kissidougou (Guinea) were *P. falciparum*-infected compared to 3/5 individuals from Uganda. To determine if the presence of *Asaia* had a quantifiable effect on the level of *P. falciparum* detected, we normalized *P. falciparum* Ct values from qPCR (Supplementary Figure 2a) and compared gene ratios for *An. gambiae* s.s. mosquitoes from Guinea, with or without *Asaia* (Supplementary Figure 2b). Statistical analysis using student’s t-tests revealed no significant difference between normalized *P. falciparum* gene ratios (p = 0.51, df =59). Larger variation of Ct values was seen for *Asaia* (Supplementary Figure 2c) suggesting the bacterial densities in individual mosquitoes were more variable than *P. falciparum* parasite infection levels.

**Discussion**

Malaria transmission in Sub-Saharan Africa is highly dependent on the local *Anopheles* vector species, but the primary vector complexes recognised are *An. gambiae* s.l., *An. funestus* s.l. *An. nili* s.l. and *An. moucheti* s.l.41,65, *An. gambiae* s.s. and *An. coluzzii* sibling species are considered the most important malaria vectors in Sub-Saharan Africa and recent studies indicate that *An. coluzzii* extends further north, and closer to the coast than *An. gambiae* s.s. within West Africa46. In our study, high *Plasmodium* prevalence rates in *An. gambiae* s.s. across Guinea would be consistent with high malaria parasite prevalence in humans (measured by rapid diagnostic tests) in Guéckédou prefecture, and the overall national malaria prevalence estimated to be 44% in 201347. However, malaria prevalence has decreased in the past few years with an overall prevalence across Guinea estimated at 15% for 2016. Although our *P. falciparum* infection prevalence rates were also high in DRC, recent studies have shown comparable levels of infection with 35% of *An. gambiae* s.l. mosquitoes infected from Kinshasa48. We detected *P. falciparum* in *An. gambiae* s.s., *An. arabiensis*, *An. funestus* s.s. and *An. species A* from DRC. Morphological differences have been widely used for identification of malaria vectors but species complexes (such as *An. gambiae* s.l. and *An. funestus* s.l.) require species-diagnostic PCR assays. Historically, malaria parasite entomology studies in Africa have focused predominantly on species from these complexes, likely due to the fact that mosquitoes from these complexes dominate the collections49. In our study, we used ITS2 sequencing to confirm secondary vector species that were *P. falciparum*-infected given the difficulties of morphological identification and recent studies demonstrating the inaccuracy of diagnostic species PCR-based molecular identification49. Our study is the first to report the detection of *P. falciparum* in *An. rufripes* from Madagascar; previously this species was considered a vector of *Plasmodium* species of non-human origin and has only very recently been implicated in human malaria transmission50. However, detection of *P. falciparum* parasites in whole body mosquitoes does not confirm that the species plays a significant role in transmission. Detection could represent infected bloodmeal stages or oocysts present in the midgut wall so further studies are warranted to determine this species ability to transmit human malaria parasites.

The mosquito microbiota can modulate the mosquito immune response and bacteria present in wild *Anopheles* populations can influence malaria vector competence42. Endosymbiotic *Wolbachia* bacteria are particularly widespread through insect populations, but they were commonly thought to be absent from *Anopheles* mosquitoes. However, the recent discovery of *Wolbachia* strains in the *An. gambiae* s.l. complex in Burkina Faso and Mali43,44, in addition to our study showing infection in *Anopheles* from Ghana and DRC, suggest resident strains could be widespread across Sub-Saharan Africa. The discovery of resident strains in Burkina Faso resulted from sequencing of the 16S rRNA gene identifying *Wolbachia*

Page 14 of 22
Figure 7. The relative abundance of bacteria in *An. gambiae* s.s. comparing two locations with contrasting *Asaia* infection densities. Bacterial genus level taxonomy was assigned to operational taxonomic units clustered with a 97% cut-off using the SILVA SSU v128 97% database, and individual genera comprising less than 1% of total abundance was merged into “Others”.

sequences rather than screening using *Wolbachia*-specific genes\(^3^9\). Intriguingly, *Wolbachia* infections in these mosquitoes could not be detected using conventional PCR targeting the *wsp* gene. As the *wsp* gene has often been used in previous studies to detect strains in *Anopheles* species\(^2^5^,2^7\), this could explain why resident strains in the *An. gambiae* s.l. complex have gone undetected until very recently. Recent similar methods using 16S rRNA amplicon sequencing to determine the overall microbiota in wild mosquito populations has provided evidence for *Wolbachia* infections in *An. gambiae* in additional villages in Burkina Faso\(^7^1\) and *Anopheles* species collected in Illinois, USA\(^7^2\). Our study describing resident *Wolbachia* strains in numerous species of *Anopheles* malaria vectors also highlights the potential for *Wolbachia* to be influencing malaria transmission, as postulated by previous studies\(^3^9^,4^0^,7^3\). Although no significant correlation was present for malaria and *Wolbachia* prevalence in the 128 *An. gambiae* s.s. individuals from DRC, we only detected co-infections in two individuals compared to 13 and 15 individuals infected only with *Wolbachia* or *P. falciparum*, respectively. As the majority (77%) of samples had neither detectable *Wolbachia* resident strains or *P. falciparum*, a larger sample size would be needed to determine if there is a correlation, as shown previously in both Burkina Faso\(^7^1\) and Mali\(^7^2\). The infection prevalence of resident *Wolbachia* strains in *An. coluzzii* from Ghana (4%) and *An. gambiae* s.s. from the DRC was variable but low (8–24%), consistent with infection prevalence in Burkina Faso (11%)\(^3^9\) but much lower than those reported in Mali (60–80%)\(^7^2\) where infection was associated with reduced prevalence and intensity of sporozoite infection in field-collected females.

The discovery of a resident *Wolbachia* strain in *An. moucheti*, a highly anthropophilic and efficient malaria vector found in the forested areas of Western and Central Africa\(^4^1\), suggests further studies are warranted that utilize large sample sizes to examine the influence of the wAnM *Wolbachia* strain on *Plasmodium* infection dynamics in this malaria vector. *An. moucheti* is often the most abundant vectors, breeding in slow moving streams and rivers, contributing to year round malaria transmission in these regions\(^7^4^,7^5\). This species has also been implicated as a main bridge vector species in the transmission of ape *Plasmodium* malaria in Gabon\(^7^6\). There is thought to be high genetic diversity in *An. moucheti* populations\(^7^7^,7^8\), which may either influence the prevalence of *Wolbachia* resident strains or *Wolbachia* could be contributing to genetic diversity through its effect on host reproduction. A novel *Wolbachia* strain in *An. species ‘A’, present at high infection frequencies in Lwiro (close to Katana in DRC), also suggests more *Anopheles* species, including unidentified and potentially new species, could be infected with this widespread endosymbiotic bacterium. *An. species A* should be further investigated to determine if this species is a potential malaria vector, given our study demonstrated
P. falciparum infection in one of two individuals screened and ELISA-positive samples of this species were reported from the Western Highlands of Kenya\(^{12}\).

The variability of Wolbachia prevalence rates in An. gambiae s.l. complex from locations within DRC and Ghana and previous studies in Burkina Faso\(^{39}\) and Mali\(^{40}\) suggest the environment is one factor that influences the presence or absence of resident strains. In our study we found no evidence of Wolbachia-Asaia co-infections across all countries, supporting laboratory studies that have shown these two bacterial endosymbionts demonstrate competitive exclusion in Anopheles species\(^{36,38}\). We also found that Asaia infection densities (whole body mosquitoes) were variable and location dependent which would correlate with this bacterium being environmentally acquired at all life stages, but also having the potential for both vertical and horizontal transmission\(^{37}\). Significant variations in overall Asaia prevalence and density across different Anopheles species and locations in our study would also correlate with our data indicating no evidence of an association with P. falciparum prevalence in both Guinea and Uganda populations. Further studies are needed to determine the complex interaction between these two bacterial endosymbionts and malaria in diverse Anopheles malaria vector species. Horizontal transfer of Wolbachia strains between species (even over large phylogenetic differences) has shaped the evolutionary history of this endosymbiont in insects, and there is evidence for loss of infection in host lineages over evolutionary time\(^{39}\). Our results showing a new strain present in An. coluzzii from Ghana (phylogenetically different to strains present in An. gambiae s.l. mosquitoes from both Burkina Faso and Mali), strain variants observed in An. species A, and the concatenated grouping of the novel Anopheles strains with strains found in different Orders of insects, support the lack of congruence between insect host and Wolbachia phylogenetic trees\(^{40}\).

Our qPCR and 16S microbiome analysis indicates the densities of wAnM and wAnsA strains are significantly higher than resident Wolbachia strains in An. gambiae s.l. However, caution must be taken as we were only able to analyse selected individuals, and larger collections of wild populations would be required to confirm these results. Native Wolbachia strains dominating the microbiome of An. species A and An. moucheti is consistent with other studies of resident strains in mosquitoes showing Wolbachia 16S rRNA gene amplicons vastly out-number sequences from other bacteria in Ae. albopictus and Cx. quinquefasciatus\(^{41,42}\). The discovery of novel Wolbachia strains provides the rationale to undertake vector competence experiments to determine what effect these strains are having on malaria transmission. The tissue tropism of novel Wolbachia strains in malaria vectors will be particularly important to characterise given this will determine if these endosymbiotic bacteria are proximal to malaria parasites within the mosquito. It would also be important to determine the additional phenotypic effects novel resident Wolbachia strains have on their mosquito hosts. Some Wolbachia strains induce a reproductive phenotype termed cytoplasmic incompatibility (CI) that results in inviable offspring when an uninfected female mates with a Wolbachia-infected male. In contrast, Wolbachia-infected females produce viable progeny when they mate with both infected and uninfected male mosquitoes. This reproductive advantage over uninfected females allows Wolbachia to spread within mosquito populations.

**Conclusions**

Wolbachia has been the focus of recent biocontrol strategies in which Wolbachia strains transferred into naïve mosquito species provide strong inhibitory effects on arboviruses\(^{16,18-20,37,44}\) and malaria parasites\(^{35,38}\). The discovery of two novel Wolbachia strains in Anopheles mosquitoes, potentially present at much higher density than resident strains in the An. gambiae s.l. complex, also suggests the potential for these strains to be transinfected into other Anopheles species to produce inhibitory effects on Plasmodium parasites. Wolbachia transinfection success is partly attributed to the relatedness of donor and recipient host so the transfer of high density Wolbachia strains between Anopheles species may result in stable infections (or co-infections) that have strong inhibitory effects on Plasmodium development. Finally, if the resident strain present in An. moucheti is at low infection frequencies in wild populations, an alternative strategy known as the incompatible insect technique (IIT) could be implemented where Wolbachia-infected males are released to suppress the wild populations through CI (reviewed by 22). In summary, the important discovery of diverse novel Wolbachia strains in Anopheles species will help our understanding of how Wolbachia strains can potentially impact malaria transmission, through natural associations or being used as candidate strains for transinfection to create stable infections in other species.

**Data availability**

16S GenBank accession numbers are listed in Supplementary Table 1; Wolbachia 16S and wsp gene GenBank accession numbers are listed in Supplementary Table 2; Wolbachia MLST gene GenBank accession numbers are listed in Supplementary Table 3.

Raw PCR screening data is available at Open Science Framework: DOI: https://doi.org/10.17605/OSF.IO/MW6XZ\(^{45}\).

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

**Grant information**

CLJ and TW were supported by a Wellcome Trust /Royal Society grant awarded to TW (101285); http://www.wellcome.ac.uk; https://royalsociety.org. GLH is supported by NIH grants (R21AI124452 and R21AI129507), a University of Texas Rising Star award, the John S. Dunn Foundation Collaborative Research Award, the Robert J. Kleberg, Jr. and Helen C. Kleberg Foundation, and the Centers for Disease Control and Prevention (CDC) (Cooperative Agreement Number U01CK000512). The papers contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Department of Health and Human Services. This work...
was also supported by a James W. McLaughlin postdoctoral fellowship at the University of Texas Medical Branch to SH. Field work in Uganda was funded by UK aid (through the Programme Partnership Arrangement grant to Malaria Consortium). YAA and ARM were supported by a NIH grant R01AI123074. SRI was funded by the U.S. President’s Malaria Initiative.

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

Acknowledgements
We would like to thank all the mosquito collectors and residents of the villages where collections took place. We would also like to thank John Gimnig, Bill Hawley and Barb Marston for reviewing our manuscript. This publication made use of the PubMLST website (https://pubmlst.org/wolbachia/) sited at the University of Oxford (Jolley & Maiden 2010, BMC Bioinformatics, 11:595). The development and maintenance of this site has been funded by the Wellcome Trust.

Supplementary material
Supplementary Table 1. Additional sample details and ITS2 GenBank accession numbers.
Click here to access the data.

Supplementary Table 2. Wolbachia 16S and wsp GenBank accession numbers.
Click here to access the data.

Supplementary Table 3. Wolbachia MLST gene GenBank accession numbers.
Click here to access the data.

Supplementary Figure 1. Alpha and beta diversity of An. gambiae s.s. from Kissidougou, Guinea and Butemba, Uganda.
Click here to access the data.

Supplementary Figure 2. Prevalence of the bacterial endosymbiont Asaia and malaria parasites in An. gambiae s.s. mosquitoes from Guinea.
Click here to access the data.

References


Open Peer Review

Current Referee Status: ?

Mathilde Gendrin Mathilde Gendrin, Ottavia Romoli
1 Microbiota of Insect Vectors Group, Institut Pasteur de la Guyane, Cayenne, French Guiana
2 Parasites and Vectors Department, Institut Pasteur, Paris, France
3 Microbiota of Insect Vectors Group, Institut Pasteur de la Guyane, Cayenne, French Guiana

Jeffries and co-authors performed a large scale analysis of the presence of Wolbachia in Anopheles mosquitoes from five countries in Africa. They found that in two of these countries, some mosquitoes were infected with Wolbachia, confirming and widening the recent discovery of the presence of Wolbachia in Anopheles mosquitoes. This is the strongest point of the paper, as an independent confirmation is always welcome and as some populations of Anopheles are even found here to have a high prevalence of Wolbachia.

The authors also checked for the presence of Asaia sp. in the analysed mosquitoes, as this bacterium is thought to compete with Wolbachia in Anopheles. They did not find any mosquito co-infected by Asaia and Wolbachia. This is also an important finding as it corroborates studies performed in the laboratory, but this time with field-collected mosquitoes. They found that in mosquitoes coming from one population, Asaia was actually a dominant species, >99% of the microbiota. Figure 7a is not very clear as one expects the scale to go from 0 to 100%, therefore we suggest to use a discontinued axis to present these interesting results.

Finally the authors investigated the presence of Plasmodium in the studied mosquitoes, as Wolbachia is thought to interfere with some transmitted pathogens. This part is less convincing as the tests have been performed on DNA extraction from whole bodies or abdomens, while the presence of Plasmodium in head and thorax (or more specifically, in salivary glands) is a more suitable method to assess transmission potential. Moreover, the conclusions drawn on the interactions between Plasmodium and Wolbachia are not exactly clear. Considering that 10.16 + 1.56 = 11.72% mosquitoes are infected with Wolbachia and 11.72 + 1.56 = 13.28% are infected with Plasmodium, if there is no effect between Wolbachia and Plasmodium, you expect that 11.72% x 13.28% = 1.56% is infected by both. Surprisingly, this is exactly the result here. Biology is rarely so close to math, for so small numbers... The authors should thus state more clearly that their results favor no interactions, as confirmed by the p value which is very close to 1. On the contrary, the discussion currently suggests that the non significant correlation is due to small numbers. However, one cannot jump to conclusion on the inability of Wolbachia to interfere with Plasmodium, as these results have been performed on abdomens and whole bodies, therefore we do not know whether the co-infected mosquitoes had just blood fed (and/or carried early stages of Plasmodium).
To improve the clarity of the article, it would be interesting to have an additional figure or table summarizing the experimental set up, explaining which mosquitoes are included in which analysis and which Wolbachia strain is found in which mosquitoes.

We also have minor comments on the manuscript:

The expression « resident strain » is not clear to us.
16S « rRNA » and « rDNA »: a consistent word may be used, rRNA seems more consensual.
The total number of mosquitoes, of Wolbachia infected mosquitoes, of Asaia infected ones, etc would be interesting.

Page 3:
§2: Asaia is not an endosymbiont
§3: « have » needs probably to be removed in « than would have been identified using morphological identification alone »
§4 needs a first sentence identifying the gap of knowledge that the authors want to fill
§5: Can the authors clearly state whether some mosquitoes had blood in their midgut?

Page 4:
Figure 1: scale should be in km, miles is not an SI unit
§2: « DNA extraction and MOSQUITO species identification ». More generally, it is not always clear whether the authors speak about mosquitoes or Wolbachia strains.
§3: « as preliminary trials revealed this was the optimal method for both sensitivity and specificity »: please add « data not shown » or remove it

Page 5:
Instead of µL of DNA, the actual quantity in ng would be preferable.
All PCRs: primer sequences are needed
§3: « Both chromatograms (forward and reverse traces) from each sample WERE manually »

Pages 6-7
Table 1: probably some mistakes, e.g. An. gambiae in Mikalayi: 11.8% corresponds neither to 1/16 nor to 2/16, so all the numbers should be checked. It would be appropriate to enter the actual numbers in brackets, and to indicate the co-prevalence of Wolbachia and Plasmodium. The legend should be grouped below or above the table and the explanation about mosquitoes in bold is unclear.
In the text, numbers would be interesting rather than only proportions.
« previously named M molecular form OF AN. GAMBAE » (or remove it, as this precision may now be superfluous). On the contrary, « An. species A » is barely introduced, it would be interesting to mention something about this species and its identification (besides the quick explanation in the introduction).

Page 13
« Approximately 1000-fold higher », it is very much of an approximation (variable Ct values and potential variations in 16S copy number): it may be good to rephrase, mentioning that 1000 is an order of magnitude rather than approximately.
§2: « An. moucheti (wAnM-infected) » comes at the 2nd occurrence of wAnM.

Is the work clearly and accurately presented and does it cite the current literature?
Yes
Is the study design appropriate and is the work technically sound?
Yes

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

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

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

Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

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