High density Novel *Wolbachia* strains in *Anopheles* species from Guinea

Claire L Jeffries[1], Cintia Cansado-Utrilla[1], Caleb Stica[1] & Thomas Walker[1]*

*corresponding author: thomas.walker@lshtm.ac.uk

[1]Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK.

Author email addresses:

Claire.Jeffries@lshtm.ac.uk

cintia.cansado-utrilla@lstmed.ac.uk

sticacaleb@gmail.com

Thomas.walker@lshtm.ac.uk

Abstract

Background: *Wolbachia*, a widespread bacterium that can influence mosquito-borne pathogen transmission, has recently been shown to infect *Anopheles* (*An*.) species that are malaria vectors in Sub-Saharan Africa. Although there are studies reporting strains in the *An. gambiae* complex, the apparent low density and low prevalence rates requires further investigation. In contrast, strains in other species appear higher density allowing a greater understanding of phylogenetics and more accurate determination of prevalence rates in wild mosquito populations.

Methods: *Anopheles* mosquitoes were collected in the Faranah and Maferinyah regions of Guinea in June-July 2018. RNA was extracted from 542 females and reverse transcribed to determine *Wolbachia* prevalence rates, demonstrate gene expression and estimate relative strain densities using quantitative PCR. Molecular confirmation of mosquito species and *Wolbachia* multilocus sequence typing (MLST) was carried out to analyse phylogenetic relationships of newly discovered strains.

Results: Low prevalence rates were detected in *An. gambiae* s.s. (0.0: 2.9%) and novel *Wolbachia* strains were discovered in two species: the *w*AnMe strain in *An. melas* (18/168: prevalence rate of 10.7%) and the *w*AnsX strain in a previously unidentified *Anopheles* species we have termed '*An.* species X' (1/1). Novel strains were phylogenetically diverse, with *w*AnMe clustering with *Wolbachia* Supergroup A strains, and *w*AnsX closest to Supergroup B strains. Significantly higher density strains were present in *An.* species X and *An. melas* compared to *An. gambiae* s.s. and were comparable in density to a novel *Wolbachia* strain discovered in *Culex watti* termed *w*Wat.

Conclusions: The discovery of novel *Wolbachia* strains provides further insight into the strain diversity within *Anopheles* species. Confirmation of gene expression for all MLST and *wsp* genes for *w*AnsX and *w*AnMe strains and qPCR analysis indicates these are candidate strains for transinfection to create stable infections in other *Anopheles* species. They have the potential for use in *Anopheles* biocontrol strategies either through population replacement or suppression control strategies.

Keywords: *Wolbachia*, mosquitoes, malaria, *Anopheles*, endosymbionts

Background

Wolbachia endosymbiotic bacteria are estimated to infect ~40% of insect species [1] and natural infections have been shown to have inhibitory effects on human arboviruses in mosquitoes [2–4]. High density *Wolbachia* strains have been utilised for mosquito biocontrol strategies targeting arboviruses as they induce synergistic phenotypic effects. *Drosophila Wolbachia* strains that have been transinfected into *Aedes (Ae.) aegypti* and *Ae. albopictus* induce inhibitory effects on arboviruses, with maternal transmission and cytoplasmic incompatibility (CI) enabling introduced strains to spread through populations [5–12]. The successful release and establishment of *Wolbachia*-transinfected *Ae. aegypti* populations in Cairns, Australia [13] was followed by further evidence of strong inhibitory effects on arboviruses from field populations [14]. Further follow up studies in Townsville, Australia [15] and Kuala Lumpur, Malaysia [16] have now shown that *Wolbachia* frequencies have remained stable since initial releases and there is a reduction in human dengue incidence in the release sites.

Wolbachia has also been postulated to have the potential to be used for biocontrol strategies targeting malaria transmission by *Anopheles* species [17]) and initial laboratory experiments were able to demonstrate that transient infections in *An. gambiae* reduce the density of *Plasmodium (P.) falciparum* parasites [18]. However, as with arboviruses there is variability in the level of inhibition of malaria parasites for different *Wolbachia* strains in different mosquito species [19–21]. A major step forward was achieved through the transinfection of a *Wolbachia* strain from *Ae. albopictus* (*w*AlbB) into *An. stephensi* and the confirmation of *P. falciparum* inhibition [22].

The interest in using *Wolbachia* for biocontrol strategies targeting malaria transmission in *Anopheles* mosquitoes further increased due to the discovery of natural strains in numerous malaria vectors of Sub-Saharan Africa [23–28]. The *An. gambiae* s.l. complex, which consists of multiple morphologically indistinguishable species, including several major malaria vector species, appears to contain diverse *Wolbachia* strains (collectively named *w*Anga) at both low prevalence in mosquito populations and at low infection densities [23,26–29]. In contrast, the recently discovered *w*AnM and *w*AnsA strains, found in *An. moucheti*

and *An.* species A respectively, are higher density infections that dominate the mosquito microbiome [24]. Interestingly, the presence of *Wolbachia* strains in *Anopheles* was inversely correlated to other bacteria species, such as *Asaia* that are stably associated with several species [30–32]. Evidence for this 'mutual exclusion' between competing bacterial species in *Anopheles* was also present from analysis of field collected mosquitoes from multiple countries in Sub-Saharan Africa [24]. In this study, we collected wild *Anopheles* mosquitoes from two regions of Guinea in June-July 2018 and characterised the natural *Wolbachia* strains to provide further evidence for the presence of these endosymbionts in malaria vectors.

Methods.

Study sites & collection methods. *Anopheles* adult mosquitoes were collected from two regions in Guinea. Human landing catches (HLCs) and larval dipping were conducted in three villages in the Faranah Prefecture; Balayani (10.1325, -10.7443), Foulaya (10.144633, -10.749717), and Tindo (9.9612230, -10.7016560) [33]. Three districts were selected for mosquito collections in the Maferinyah sub-prefecture using a variety of traps [34]. BG sentinel 2 traps (BG2) (Biogents), CDC light traps (John W. Hock), gravid traps (BioQuip) and stealth traps (John W. Hock) were used to sample adult mosquitoes in Maferinyah Centre (09.54650, -013.28160), Senguelen (09.41150, -013.37564) and Fandie **(**09.53047, -013.24000). Mosquitoes collected from traps and HLCs were morphologically identified using keys and stored in RNAlater® (Invitrogen) at -80ºC [33,34].

RNA extraction and generation of cDNA. RNA was extracted from individual whole mosquitoes using Qiagen 96 RNeasy Kits according to manufacturer's instructions and a Qiagen Tissue Lyser II (Hilden, Germany) with a 5mm stainless steel bead (Qiagen) per sample to homogenise mosquitoes. RNA was eluted in 45 µL of RNase-free water and stored at -70˚C. RNA was reverse transcribed into complementary DNA (cDNA) using an Applied Biosystems High Capacity cDNA Reverse Transcription kit. A final volume of 20 µL contained 10 µL RNA, 2 µL 10X RT buffer, 0.8 µL 25X dNTP (100 mM), 2 µL 10X random primers, 1µL reverse transcriptase and 4.2 µL nuclease-free water. Reverse transcription was undertaken

in a Bio-Rad T100 Thermal Cycler as follows: 25ºC for 10min, 37ºC for 120min and 85ºC for 5min and cDNA stored at –20°C.

Molecular mosquito species identification. Species identification of the *An. gambiae* complex was initially undertaken using diagnostic species-specific PCR assays to distinguish between the morphologically indistinguishable sibling species [35,36]. To determine species identification for samples of interest and samples that could not be identified by species-specific PCR, sequences were generated by Sanger sequencing of *ITS2* PCR products [34,37].

***Wolbachia* detection and amplification of *Wolbachia* genes.** *Wolbachia* detection was first undertaken targeting the conserved *Wolbachia* genes previously shown to amplify a wide diversity of strains; *16S rDNA* gene using primers W-Spec-16S-F: 5'-CATACCTATTCGAAGGGATA-3' and W-Spec-16s-R: 5'-AGCTTCGAGTGAAACCAATTC-3 [38] and *Wolbachia* surface protein (*wsp*) gene using primers wsp81F: 5'-TGGTCCAATAAGTGATGAAGAAAC-3' and wsp691R: 5'-AAAAATTAAACGCTACTCCA-3' [39]. PCR analysis was also undertaken on cDNA to determine any evidence for the presence of CI-inducing genes *CifA* (primers 5'-TGTGGTAGGGAAGGAAAGAGGAAA-3', 5'-ATTCCAAGGACCATCACCTACAGA-3') and *CifB* (primers 5'-TGCGAGAGATTAGAGGGCAAAATC-3', 5'-CCTAAGAAGGCTAATCTCAGACGC-3') [40]. Multilocus strain typing (MLST) was undertaken to characterize *Wolbachia* strains using the sequences of five conserved genes as molecular markers to genotype each strain. In brief, 450-500 base pair fragments of the *gatB*, *coxA*, *hcpA*, *ftsZ* and *fbpA Wolbachia* genes were amplified from individual *Wolbachia*-infected mosquitoes using previously optimised protocols [41]. Primers used were as follows: gatB_F1: 5'-GAKTTAAAYCGYGCAGGBGTT-3', gatB_R1: 5'-TGGYAAYTCRGGYAAAGATGA-3', coxA_F1: 5'-TTGGRGCRATYAACTTTATAG-3', coxA_R1: 5'-CTAAAGACTTTKACRCCAGT-3', hcpA_F1: 5'-GAAATARCAGTTGCTGCAAA-3', hcpA_R1: 5'-GAAAGTYRAGCAAGYTCTG-3', ftsZ_F1: 5'-ATYATGGARCATATAAARGATAG-3', ftsZ_R1: 5'-TCRAGYAATGGATTRGATAT-3', fbpA_F1: 5'-GCTGCTCCRCTTGGYWTGAT-3' and fbpA_R1: 5'-CCRCCAGARAAAAYYACTATTC-3'. If initial amplification with

these primers was unsuccessful, the PCR was repeated using the standard primers but with the addition of M13 adaptors. If no amplification was detected using standard primers, further PCR analysis was undertaken using degenerate primer sets, with or without M13 adaptors [41]. 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.

Sanger sequencing. 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. Where *Wolbachia* PCR primers included M13 adaptors, just the M13 primers alone (M13_adaptor_F: 5'-TGTAAAACGACGGCCAGT-3' and M13_adaptor_R: 5'-CAGGAAACAGCTATGACC-3') were used for sequencing, otherwise the same primers as utilised for PCR were used. Sequencing analysis was carried out in MEGA7 [42]. Both chromatograms (forward and reverse traces) from each sample were manually checked, edited, and trimmed as required, followed by alignment by ClustalW and checking to produce consensus sequences. Consensus sequences were used to perform nucleotide BLAST (NCBI) database queries, and for *Wolbachia* genes searches against the *Wolbachia* MLST database (http://pubmlst.org/wolbachia). 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 for *Anopheles Wolbachia* strains 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 [43]. 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 automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. The phylogeny test was by Bootstrap method with 1000 replications. Evolutionary analyses were conducted in MEGA7 [42].

***Wolbachia* quantification.** To estimate *Wolbachia* density across multiple mosquito species, RNA extracts were added to Qubit™ RNA High Sensitivity Assays (Invitrogen) and total RNA was measured using a Qubit 4 Fluorometer (Invitrogen). All RNA extracts were then diluted to produce extracts that were 2.0 nanograms (ng)/µL prior to being used in quantitative Reverse Transcription PCR (qRT-PCR) assays targeting the *16S rRNA* gene [26]. A synthetic oligonucleotide standard (Integrated DNA Technologies) was designed to calculate *16S rDNA* gene copies per µL using a ten-fold serial dilution (**Supplementary figure 1**). *16S* rDNA gene real time qRT-PCR reactions were prepared using 5 µL of QuantiNova SYBR Green RT-PCR Kit (Qiagen), a final concentration of 1µM 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. Each mosquito RNA extract was run in triplicate alongside standard curves and NTCs and PCR results were analysed using the LightCycler® 96 software (Roche Diagnostics).

Asaia detection. *Asaia* PCR screening was undertaken by targeting the 16S *rRNA* gene using primers Asafor: 5’-GCGCGTAGGCGGTTTACAC-3’ and

Asarev: 5'-AGCGTCAGTAATGAGCCAGGTT-3' [31,44]. *Asaia 16S* rDNA gene real time qRT-PCR reactions were prepared using 5 µL of QuantiNova SYBR Green RT-PCR Kit, a final concentration of 1µM 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.

Statistical analysis. Normalised qRT-PCR *Wolbachia 16S* rRNA gene copies per µL were compared using unpaired t-tests in GraphPad Prism 7.

Results

Mosquito species and *Wolbachia* strain prevalence rates. Prevalence rates of natural *Wolbachia* strains were variable depending on *Anopheles* species and location (**Table 1**). *Wolbachia* strains were detected in *An. gambiae* s.s. mosquitoes from Faranah with prevalence rates ranging from 0.0 - 2.9% (termed *w*Anga-Guinea) and *Wolbachia* in *An. melas* (10.7% prevalence – termed *w*AnMe) and in the only female unidentified *Anopheles* species 'X' (termed *w*AnsX) from Senguelen in the Maferinyah sub-prefecture. The molecular phylogeny of the *ITS2* gene confirmed molecular mosquito species identification of the unidentified *Anopheles* species 'X' (**Figure 1**) and *An. melas* (**Figure 2**). All *ITS2* sequences were deposited in GenBank (accession numbers XXXXX – XXXXX) (**Supplementary Table 1)**. The *ITS2* fragment sequenced from the unidentified *Anopheles* species 'X' was most similar to *Anopheles* sp. 7 BSL-2014 (GenBank accession number KJ522819.1) but at only 93.2% sequence identity, and *An. theileri* (GenBank accession number MH378771.1) with 90.9% sequence identity (both full query coverage).

***Wolbachia* strain typing**. Although we were able to amplify the *16S rRNA* fragments of the natural strain in *An. gambiae* s.s., we were unable to obtain sequences of sufficient quality and were also unable to amplify the *wsp* gene. In contrast, we obtained *16S rRNA* (**Figure 3**) and *wsp* sequences (**Figure 4**) from both *w*AnsX and *w*AnMe strains. Phylogenetic analysis of both genes shows that the *w*AnsX strain is most closely related to *Wolbachia* strains of Supergroup B (such as

*w*Pip, *w*AlbB, *w*AnsA, *w*AnM, *w*Ma and *w*No). In contrast, the *w*AnMe strain is most closely related to *Wolbachia* strains of Supergroup A (such as *w*Mel, *w*AlbA and *w*Au). Typing of the *w*AnsX *wsp* nucleotide sequence highlighted that there were no exact matches to *wsp* alleles currently in the *Wolbachia* MLST database (https://pubmlst.org/wolbachia/), and only one of the four hypervariable regions (HVRs) matched a known sequence (HVR3: allele 3). Whereas the *w*AnMe *wsp* sequence matched allele 23 within the database (**Table 2**). All *Wolbachia* gene sequences of sufficient quality to generate a consensus were deposited into GenBank (accession numbers XXXXX – XXXXX) (**Supplementary Tables 2 and 3**).

MLST was undertaken to provide more accurate strain phylogenies. This was successfully done for the novel *Wolbachia* strains *w*AnsX and *w*AnMe but we were unable to amplify any of the five MLST genes from *Wolbachia*-infected *An. gambiae* s.s. from Faranah. New alleles for all five MLST gene loci (sequences differed from those currently present in the MLST database) for wAnsX, and novel allelic profiles for both strains confirm the diversity of these novel *Wolbachia* strains (**Table 2**). The phylogeny of *w*AnsX based on concatenated sequences of all five MLST gene loci confirms this strain clusters within Supergroup B (**Figure 5).** This demonstrates this is a novel strain 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 5, Table 2**). MLST gene fragment amplification was more variable for *w*AnMe requiring *hcpA* 'A strain' specific primers (hcpA_F1: GAAATARCAGTTGCTGCAAA, hcpA_AspecR1: TTCTARYTCTTCAACCAATGC) to generate sequence of sufficient quality for analysis of the hcpA gene (hcpA_F1). Concatenation of the MLST loci also confirms *w*AnMe is closest to strains belonging to Supergroup A, including *w*Mel and *w*AlbA (as suggested by *16S* and *wsp* gene phylogenies). Consistent with our previous study looking at novel *Wolbachia* strains in *Anopheles* species using MLST [24], these results highlight the lack of concordance between *Wolbachia* strain phylogeny and their insect hosts across diverse geographical regions. Our PCR analysis also did not result in any amplification of *CifA* or *CifB* gene fragments for either the *w*AnsX or *w*AnMe strains, although we are unable to determine if this could be due to sequence variation preventing amplification given these are novel *Wolbachia* strains.

***Wolbachia* strain densities and relative abundance**. The relative densities of *Wolbachia* strains were estimated using qRT-PCR targeting the *16S rRNA* gene after first standardising total RNA (ng per reaction). This allowed direct comparisons between phylogenetically diverse *Anopheles* species and accounts for variation in mosquito body size and DNA extraction efficiency between samples. This also allows a comparison to another novel natural *Wolbachia* strain present in *Cx. watti* (termed *w*Wat strain) collected in Maferinyah, the same sub-prefecture of Guinea. *w*Wat strain gene sequences were deposited into GenBank (accession numbers XXXXX – XXXXX). *16S rRNA* qRT-PCR analysis revealed a mean of 1.50E+04 (+/- 4.37E+03) *16S rRNA* copies/µL for the *w*AnsX strain in the single individual **(Figure 6, Supplementary table 4).** A lower mean density was found for the *w*AnMe strain in *An. melas* individuals (n=18) with 6.69E+02 (+/- 2.34E+02) *16S rRNA* copies/µL. We compared the densities to the *w*Wat strain in *Cx. watti* females also collected in the Maferinyah region and found a mean density of 2.37E+04 (+/- 5.99E+03). The density of the *w*Wat strain was significantly higher than the *w*AnMe strain (Unpaired T-test, p=0.002). Individual *An. gambiae* s.s. extracts that were identified as *Wolbachia*-infected by amplification of the *16S rRNA* gene [38] did not result in any *16S rRNA* qRT-PCR amplification suggesting a very low titre *Wolbachia* strain present in this species.

***Wolbachia* and *Asaia* co-infections.** Individual mosquitoes shown to be infected with the *w*AnsX or *w*AnMe strain were screened for the presence of *Asaia* bacteria using qRT-PCR. Co-infections were detected in both the single *An. species* X (*Asaia 16S rRNA* Ct value = 34.92) and in all *An. melas* (n=18, mean *Asaia 16S rRNA* Ct value = 29.65 +/- 2.95) (**Supplementary table 4**).

Discussion

Endosymbiotic *Wolbachia* bacteria are particularly widespread through insect populations but were historically considered absent from the *Anopheles* genera [45]. The discovery of additional novel natural strains of *Wolbachia* in *Anopheles* species suggests that the prevalence and diversity has been significantly under-reported to date. Since 2014, there have been several reports of *Wolbachia* strains in major malaria vectors, such as sibling species in the *An. gambiae* complex [23,26–29] and

An. moucheti [24]. This study provides strong evidence for *Wolbachia* strains in *An. melas*, a species within the *An. gambiae* complex, which can be an important local vector of malaria in West-African coastal areas where it breeds in brackish water, mangrove forests and salt marshes [46,47]. It's importance as a local malaria vector was shown in Equatorial Guinea where the average number of malaria infective *An. melas* bites/person/year was recorded at up to 130 [48].

The discovery of the *w*AnsX strain (with a high *Wolbachia* titre measured by qRT-PCR) led to retrospective confirmation of the host mosquito species using *ITS2* Sanger sequencing. In this study and previous studies, accurate molecular identification is important given the difficulties of morphological identification and inaccuracies of diagnostic species PCR-based molecular identification [49]. Our *ITS2* analysis revealed that this species of *Anopheles,* which we have termed 'species X', is most closely related to *Anopheles* sp. 7 BSL-2014 (GenBank accession number KJ522819.1) and *An. theileri* (GenBank accession number MH378771.1). *Anopheles* sp. 7 BSL-2014 was collected in the Western Kenyan Highlands but little is known about this species [50]. *An. theileri* was collected in the Democratic Republic of Congo [51] and was found to be infected with *Plasmodium* sporozoites in eastern Zambia [52].

The results of this study also highlight the requirement to provide as much genetic information as possible for a newly discovered strain of *Wolbachia* (particularly low titre infections). The first discovery of *Wolbachia* strains in wild *An. gambiae* populations in Burkina Faso resulted from sequencing of the *16S rRNA* gene identifying *Wolbachia* sequences, rather than screening using *Wolbachia*-specific genes [28]. A more recent comprehensive analysis of the *An. gambiae* complex through screening of *An. gambiae* genomes (Ag1000G project) concluded that determining whether a *Wolbachia* strain is present in a given host based on the sequencing of one gene fragment (often *16S rRNA*) is problematic and caution should be taken [29]. In this study, we were only able to amplify a *Wolbachia 16S rRNA* gene fragment from *An. gambiae* s.s., which is consistent with numerous recent studies in which low density strains have been detected [23,25]. As a result, caution must be taken in the biological significance of this strain. Other explanations for the amplification of *16S rRNA* gene fragments include *Wolbachia* DNA insertions

into an insect chromosome or contamination from non-mosquito material such as ectoparasites or plants [29]. In contrast to previous studies, we extracted RNA from *An. gambiae* s.s. individuals, demonstrating expression of the *16S rRNA* gene, and indicating amplification is more likely of bacterial gene origin, rather than through integration into the host genome. However, these results are consistent with previous studies in which every *Wolbachia 16S rRNA* amplicon and sequence attributed to *An. gambiae* is unique and appears at very low titre [29].

The densities of the *w*AnsX and *w*AnMe strains (measured using qRT-PCR) are significantly higher than resident *Wolbachia* strains in *An. gambiae* s.s. (which were not detectable using this assay targeting the *16S rRNA* gene). However, caution is also required for the *w*AnsX strain as we only collected one individual of this unidentified *Anopheles* species. The *w*AnMe strain appears to have both an intermediate prevalence rate and density and further studies are required to investigate this strain across more diverse geographical areas. The detection of *Wolbachia-Asaia* co-infections in *An.* species X and *An. melas* was in contrast to our previous study [24] but *Asaia* can be environmentally acquired at different mosquito life stages and the prevalence and density was significantly variable across different *Anopheles* species and locations [24]. These contrasting results suggest a complex association between these two bacterial species in wild *Anopheles* mosquito populations and given that *Asaia* is environmentally acquired, this association will be highly location-dependent.

Wolbachia strains in *An. species A* (*w*AnsA) and *An. moucheti* (*w*AnM) [24] and now *An. melas* (*w*AnMe) and *An. species X* (*w*AnsX) have complete MLST and *wsp* profiles and are at significantly higher densities when compared to strains detected in *An. gambiae* s.s. from the same countries. As *Wolbachia* density is strongly correlated with arbovirus inhibition in *Aedes* mosquitoes [5,6,11,53], higher density strains in *Anopheles* species would be predicted to have a greater impact on malaria transmission in field populations. In this study, we screened for *P. falciparum* infection and found very low prevalence rates (<1%) (data not shown) preventing any statistical analysis on *Wolbachia-Plasmodium* interactions. This study and previous studies measuring a direct impact on *Plasmodium* infection in wild populations are dependent on parasite infection rates which can be low even in

malaria-endemic areas [24] and particularly for the infective sporozoite stage [54]. Low pathogen prevalence rates are also limiting factors in assessing the effect of natural strains of *Wolbachia* on arboviruses in wild mosquito populations [55]. In addition to looking at effects on *Plasmodium* prevalence in field populations, further work should look to undertake vector competence experiments with colonised populations and to determine if these *Wolbachia* strains are present in tissues such as the midgut and salivary glands which are critical to sporogony. Furthermore, an assessment of how these *Wolbachia* strains are being maintained in field populations is needed, and to determine if the CI reproductive phenotype can be induced by these strains (despite no evidence for this from our PCR analysis).

Conclusions

The discovery of two additional novel *Wolbachia* strains in *Anopheles* mosquitoes, present at higher density than resident strains in *An. gambiae* s.s. in Guinea, provides further evidence of the under-reporting of natural strains in the *Anopheles* genus. Although the debate continues over the biological significance (or even presence of natural strains in the *An. gambiae* complex), this study provides strong evidence of two additional novel strains with higher density infections. Candidate *Wolbachia* strains for mosquito biocontrol strategies require synergistic phenotypic effects to impact the transmission of mosquito-borne pathogens and further studies are needed to determine if these strains would induce CI and what effects they may have on host fitness.

Competing interests

No competing interests were disclosed.

Grant information

CLJ and TW were supported by Wellcome Trust /Royal Society grants awarded to TW (101285/Z/13/Z): http://www.wellcome.ac.uk; https://royalsociety.org.

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

Acknowledgements

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.

References

1. Zug R, Hammerstein P. Still a host of hosts for Wolbachia: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PLoS One. 2012/06/12. 2012;7: e38544. doi:10.1371/journal.pone.0038544
2. Glaser RL, Meola MA. The native Wolbachia endosymbionts of Drosophila melanogaster and Culex quinquefasciatus increase host resistance to West Nile virus infection. PLoS One. 2010/08/12. 2010;5: e11977. doi:10.1371/journal.pone.0011977
3. Mousson L, Zouache K, Arias-Goeta C, Raquin V, Mavingui P, Failloux AB. The native Wolbachia symbionts limit transmission of dengue virus in Aedes albopictus. PLoS Negl Trop Dis. 2013/01/10. 2012;6: e1989. doi:10.1371/journal.pntd.0001989
4. Silva JBL, Alves DM, Bottino-Rojas V, Pereira TN, Sorgine MHF, Caragata EP, et al. Wolbachia and dengue virus infection in the mosquito Aedes fluviatilis (Diptera: Culicidae). PLoS One. 2017;12. doi:10.1371/journal.pone.0181678
5. Walker T, Johnson PH, Moreira LA, Iturbe-Ormaetxe I, Frentiu FD, McMeniman CJ, et al. The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations. Nature. 2011; doi:10.1038/nature10355
6. Joubert DA, Walker T, Carrington LB, De Bruyne JT, Kien DHT, Hoang NLT, et al. Establishment of a Wolbachia Superinfection in Aedes aegypti Mosquitoes as a Potential Approach for Future Resistance Management. PLoS Pathog. 2016;12. doi:10.1371/journal.ppat.1005434
7. Moreira LA, Iturbe-Ormaetxe I, Jeffery JA, Lu G, Pyke AT, Hedges LM, et al. A Wolbachia symbiont in Aedes aegypti limits infection with dengue, Chikungunya, and Plasmodium. Cell. 2010/01/13. 2009;139: 1268–1278. doi:10.1016/j.cell.2009.11.042
8. Bian G, Xu Y, Lu P, Xie Y, Xi Z. The endosymbiotic bacterium Wolbachia

induces resistance to dengue virus in Aedes aegypti. PLoS Pathog. 2010/04/07. 2010;6: e1000833. doi:10.1371/journal.ppat.1000833

9. Blagrove MSC, Arias-Goeta C, Failloux A-B, Sinkins SP. Wolbachia strain wMel induces cytoplasmic incompatibility and blocks dengue transmission in Aedes albopictus. Proc Natl Acad Sci. 2012;109: 255–260. doi:10.1073/pnas.1112021108
10. Ant TH, Herd CS, Geoghegan V, Hoffmann AA, Sinkins SP. The Wolbachia strain wAu provides highly efficient virus transmission blocking in Aedes aegypti. PLoS Pathog. 2018;14. doi:10.1371/journal.ppat.1006815
11. Fraser JE, De Bruyne JT, Iturbe-Ormaetxe I, Stepnell J, Burns RL, Flores HA, et al. Novel Wolbachia-transinfected Aedes aegypti mosquitoes possess diverse fitness and vector competence phenotypes. PLoS Pathog. 2017;13. doi:10.1371/journal.ppat.1006751
12. McMeniman CJ, Lane R V, Cass BN, Fong AW, Sidhu M, Wang YF, et al. Stable introduction of a life-shortening Wolbachia infection into the mosquito Aedes aegypti. Science (80-). 2009/01/03. 2009;323: 141–144. doi:10.1126/science.1165326
13. Hoffmann AA, Montgomery BL, Popovici J, Iturbe-Ormaetxe I, Johnson PH, Muzzi F, et al. Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission. Nature. 2011/08/26. 2011;476: 454–457. doi:10.1038/nature10356
14. Frentiu FD, Zakir T, Walker T, Popovici J, Pyke AT, van den Hurk A, et al. Limited Dengue Virus Replication in Field-Collected Aedes aegypti Mosquitoes Infected with Wolbachia. PLoS Negl Trop Dis. 2014/03/04. 2014;8: e2688. doi:10.1371/journal.pntd.0002688
15. O'Neill SL, Ryan PA, Turley AP, Wilson G, Retzki K, Iturbe-Ormaetxe I, et al. Scaled deployment of Wolbachia to protect the community from dengue and other Aedes transmitted arboviruses. Gates Open Res. 2018; doi:10.12688/gatesopenres.12844.2
16. Nazni WA, Hoffmann AA, Noor Afizah A, Cheong YL, Mancini M V, Golding N, et al. Establishment of *Wolbachia* strain *w*AlbB in Malaysian populations of *Aedes aegypti* for dengue control. bioRxiv. 2019; 775965. doi:10.1101/775965
17. Walker T, Moreira LA. Can Wolbachia be used to control malaria? Mem Inst Oswaldo Cruz. 2011/09/09. 2011;106 Suppl: 212–217. Available:

http://www.ncbi.nlm.nih.gov/pubmed/21881776

18. Hughes GL, Koga R, Xue P, Fukatsu T, Rasgon JL. Wolbachia infections are virulent and inhibit the human malaria parasite Plasmodium falciparum in Anopheles gambiae. PLoS Pathog. 2011/06/01. 2011;7: e1002043. doi:10.1371/journal.ppat.1002043
19. Hughes GL, Vega-Rodriguez J, Xue P, Rasgon JL. Wolbachia strain wAlbB enhances infection by the rodent malaria parasite Plasmodium berghei in Anopheles gambiae mosquitoes. Appl Environ Microbiol. 2012; doi:10.1128/AEM.06751-11
20. Murdock CC, Blanford S, Hughes GL, Rasgon JL, Thomas MB. Temperature alters Plasmodium blocking by Wolbachia. Sci Rep. 2014; doi:10.1038/srep03932
21. Hughes GL, Rivero A, Rasgon JL. Wolbachia Can Enhance Plasmodium Infection in Mosquitoes: Implications for Malaria Control? PLoS Pathog. 2014; doi:10.1371/journal.ppat.1004182
22. Bian G, Joshi D, Dong Y, Lu P, Zhou G, Pan X, et al. Wolbachia invades Anopheles stephensi populations and induces refractoriness to Plasmodium infection. Science (80-). 2013/05/11. 2013;340: 748–751. doi:10.1126/science.1236192
23. Baldini F, Rougé J, Kreppel K, Mkandawile G, Mapua SA, Sikulu-Lord M, et al. First report of natural Wolbachia infection in the malaria mosquito Anopheles arabiensis in Tanzania. Parasites and Vectors. 2018; doi:10.1186/s13071-018-3249-y
24. Jeffries CL, Lawrence GG, Golovko G, Kristan M, Orsborne J, Spence K, et al. Novel Wolbachia strains in Anopheles malaria vectors from Sub-Saharan Africa. Wellcome Open Res. 2018;3: 113. doi:10.12688/wellcomeopenres.14765.2
25. Niang EHA, Bassene H, Makoundou P, Fenollar F, Weill M, Mediannikov O. First report of natural Wolbachia infection in wild Anopheles funestus population in Senegal. Malar J. 2018; doi:10.1186/s12936-018-2559-z
26. Gomes FM, Hixson BL, Tyner MDW, Ramirez JL, Canepa GE, Alves e Silva TL, et al. Effect of naturally occurring *Wolbachia* in *Anopheles gambiae s.l.* mosquitoes from Mali on *Plasmodium falciparum* malaria transmission. Proc Natl Acad Sci. 2017; 201716181. doi:10.1073/pnas.1716181114

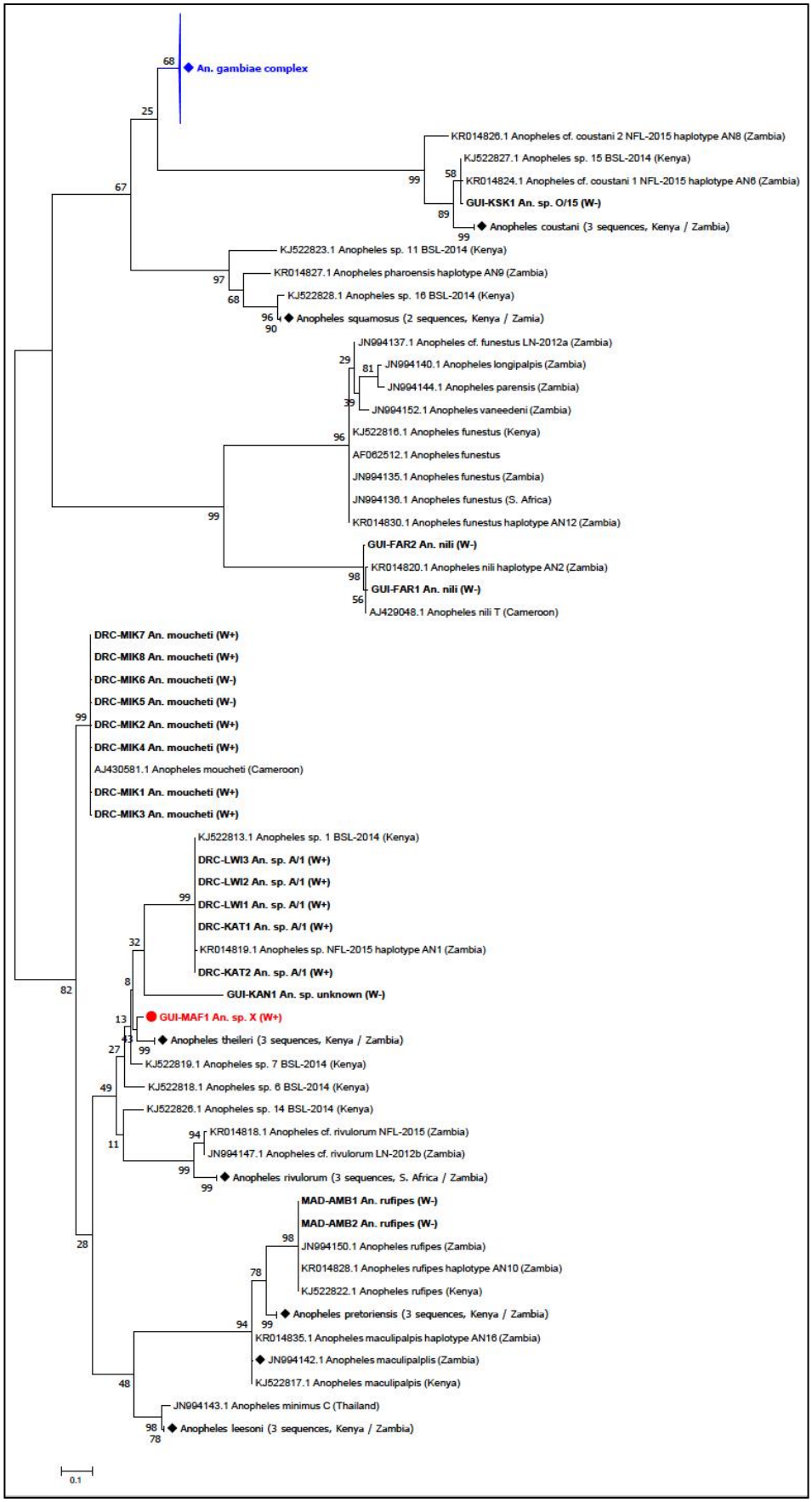
27. Shaw WR, Marcenac P, Childs LM, Buckee CO, Baldini F, Sawadogo SP, et al. Wolbachia infections in natural Anopheles populations affect egg laying and negatively correlate with Plasmodium development. Nat Commun. 2016;7. doi:10.1038/ncomms11772
28. Baldini F, Segata N, Pompon J, Marcenac P, Robert Shaw W, Dabiré RK, et al. Evidence of natural Wolbachia infections in field populations of Anopheles gambiae. Nat Commun. 2014;5. doi:10.1038/ncomms4985
29. Chrostek E, Gerth M. Is Anopheles gambiae a natural host of Wolbachia? bioRxiv. 2018; doi:10.1101/491449
30. Hughes GL, Dodson BL, Johnson RM, Murdock CC, Tsujimoto H, Suzuki Y, et al. Native microbiome impedes vertical transmission of Wolbachia in Anopheles mosquitoes. Proc Natl Acad Sci. 2014;111: 12498–12503. doi:10.1073/pnas.1408888111
31. Favia G, Ricci I, Damiani C, Raddadi N, Crotti E, Marzorati M, et al. Bacteria of the genus Asaia stably associate with Anopheles stephensi, an Asian malarial mosquito vector. Proc Natl Acad Sci U S A. 2007/05/16. 2007;104: 9047–9051. doi:10.1073/pnas.0610451104
32. Rossi P, Ricci I, Cappelli A, Damiani C, Ulissi U, Mancini MV, et al. Mutual exclusion of *Asaia* and *Wolbachia* in the reproductive organs of mosquito vectors. Parasit Vectors. 2015; doi:10.1186/s13071-015-0888-0
33. Stica C, Jeffries CL, Irish SR, Barry Y, Camara D, Yansane I, et al. Characterizing the molecular and metabolic mechanisms of insecticide resistance in Anopheles gambiae s.l. in Faranah, Guinea. bioRxiv. 2019; 610998. doi:10.1101/610998
34. Cansado Utrilla C, Jeffries CL, Kristan M, Brugman VA, Heard P, Camara G, et al. An assessment of adult mosquito collection techniques for studying species abundance and diversity in Maferinyah, Guinea. bioRxiv. 2019; 772822. doi:10.1101/772822
35. Scott JA, Brogdon WG, Collins FH. Identification of single specimens of the Anopheles gambiae complex by the polymerase chain reaction. Am J Trop Med Hyg. 1993/10/01. 1993;49: 520–529. Available: http://www.ncbi.nlm.nih.gov/pubmed/8214283
36. Santolamazza F, Mancini E, Simard F, Qi Y, Tu Z, della Torre A. Insertion polymorphisms of SINE200 retrotransposons within speciation islands of

Anopheles gambiae molecular forms. Malar J. 2008/08/30. 2008;7: 163. doi:10.1186/1475-2875-7-163

37. Beebe N, Saul A. Discrimination of all members of the Anopheles punctulatus complex by polymerase chain reaction restriction fragment length polymorphism analysis. Am J Trop Med Hyg. 1995;53: 478–481. Available: http://europepmc.org/abstract/MED/7485705
38. Werren JH, Windsor DM. Wolbachia infection frequencies in insects: evidence of a global equilibrium? Proc Biol Sci. 2000/09/06. 2000;267: 1277–1285. doi:10.1098/rspb.2000.1139
39. Zhou W, Rousset F, O'Neil S. Phylogeny and PCR-based classification of Wolbachia strains using wsp gene sequences. Proc Biol Sci. 1998/05/07. 1998;265: 509–515. doi:10.1098/rspb.1998.0324
40. Le Page DP, Metcalf JA, Bordenstein SR, On J, Perlmutter JI, Shropshire JD, et al. Prophage WO genes recapitulate and enhance Wolbachia-induced cytoplasmic incompatibility. Nature. 2017;543: 243–247. doi:10.1038/nature21391
41. Jolley KA, Chan MS, Maiden MCJ. mlstdbNet - Distributed multi-locus sequence typing (MLST) databases. BMC Bioinformatics. 2004;5. doi:10.1186/1471-2105-5-86
42. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33: 1870–1874. doi:10.1093/molbev/msw054
43. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993;10: 512–26. doi:10.1093/molbev/msl149
44. Yamada Y, Katsura K, Kawasaki H, Widyastuti Y, Saono S, Seki T, et al. Asaia bogorensis gen. nov., sp. nov., an unusual acetic acid bacterium in the alpha-Proteobacteria. Int J Syst Evol Microbiol. 2000;50 Pt 2: 823–829. doi:10.1099/00207713-50-2-823
45. Walker T, Moreira LA. Can Wolbachia be used to control malaria? Mem Inst Oswaldo Cruz. 2011;106 Suppl: 212–217. Available: http://www.ncbi.nlm.nih.gov/pubmed/21881776
46. Bryan JH, Petrarca V, Di Deco MA, Coluzzi M. Adult behaviour of members of the Anopheles gambiae complex in the Gambia with special reference to An.

melas and its chromosomal variants. Parassitologia. 1987;
47. Caputo B, Nwakanma D, Jawara M, Adiamoh M, Dia I, Konate L, et al. Anopheles gambiae complex along The Gambia river, with particular reference to the molecular forms of An. gambiae s.s. Malar J. 2008/09/23. 2008;7: 182. doi:10.1186/1475-2875-7-182
48. Overgaard HJ, Reddy VP, Abaga S, Matias A, Reddy MR, Kulkarni V, et al. Malaria transmission after five years of vector control on Bioko Island, Equatorial Guinea. Parasit Vectors. 2012/11/14. 2012;5: 253. doi:10.1186/1756-3305-5-253
49. Erlank E, Koekemoer LL, Coetzee M. The importance of morphological identification of African anopheline mosquitoes (Diptera: Culicidae) for malaria control programmes. Malaria Journal. 2018. doi:10.1186/s12936-018-2189-5
50. Laurent BS, Cooke M, Krishnankutty SM, Asih P, Mueller JD, Kahindi S, et al. Molecular characterization reveals diverse and unknown malaria vectors in the western Kenyan highlands. Am J Trop Med Hyg. 2016; doi:10.4269/ajtmh.15-0562
51. Bandibabone J, Muhigwa JBB, Agramonte NM, Zawadi B, Ombeni L, Corredor-Medina C, et al. Identification of Anopheles species in Sud Kivu, Democratic Republic of Congo, using molecular tools. Trans R Soc Trop Med Hyg. 2018; doi:10.1093/trstmh/try069
52. Lobo NF, St. Laurent B, Sikaala CH, Hamainza B, Chanda J, Chinula D, et al. Unexpected diversity of Anopheles species in Eastern Zambia: Implications for evaluating vector behavior and interventions using molecular tools. Sci Rep. 2015;5. doi:10.1038/srep17952
53. Ant TH, Herd CS, Geoghegan V, Hoffmann AA, Sinkins SP. The Wolbachia strain wAu provides highly efficient virus transmission blocking in Aedes aegypti. PLoS Pathog. 2018; doi:10.1371/journal.ppat.1006815
54. Pombi M, Calzetta M, Guelbeogo WM, Manica M, Perugini E, Pichler V, et al. Unexpectedly high Plasmodium sporozoite rate associated with low human blood index in Anopheles coluzzii from a LLIN-protected village in Burkina Faso. Sci Rep. 2018; doi:10.1038/s41598-018-31117-x
55. Jeffries CL, Tantely LM, Raharimalala FN, Hurn E, Boyer S, Walker T. Diverse novel resident Wolbachia strains in Culicine mosquitoes from Madagascar. Sci Rep. 2018; doi:10.1038/s41598-018-35658-z

Figures


An. gambiae complex
KR014826.1 Anopheles cf. coustani 2 NFL-2015 haplotype AN8 (Zambia)
KJ522827.1 Anopheles sp. 15 BSL-2014 (Kenya)
KR014824.1 Anopheles cf. coustani 1 NFL-2015 haplotype AN6 (Zambia)
GUI-KSK1 An. sp. O/15 (W-)
Anopheles coustani (3 sequences, Kenya / Zambia)
KJ522823.1 Anopheles sp. 11 BSL-2014 (Kenya)
KR014827.1 Anopheles pharoensis haplotype AN9 (Zambia)
KJ522828.1 Anopheles sp. 16 BSL-2014 (Kenya)
Anopheles squamosus (2 sequences, Kenya / Zamia)
JN994137.1 Anopheles cf. funestus LN-2012a (Zambia)
JN994140.1 Anopheles longipalpis (Zambia)
JN994144.1 Anopheles parensis (Zambia)
JN994152.1 Anopheles vaneedeni (Zambia)
KJ522816.1 Anopheles funestus (Kenya)
AF062512.1 Anopheles funestus
JN994135.1 Anopheles funestus (Zambia)
JN994136.1 Anopheles funestus (S. Africa)
KR014830.1 Anopheles funestus haplotype AN12 (Zambia)
GUI-FAR2 An. nili (W-)
KR014820.1 Anopheles nili haplotype AN2 (Zambia)
GUI-FAR1 An. nili (W-)
AJ429048.1 Anopheles nili T (Cameroon)
DRC-MIK7 An. moucheti (W+)
DRC-MIK8 An. moucheti (W+)
DRC-MIK6 An. moucheti (W-)
DRC-MIK5 An. moucheti (W-)
DRC-MIK2 An. moucheti (W+)
DRC-MIK4 An. moucheti (W+)
AJ430581.1 Anopheles moucheti (Cameroon)
DRC-MIK1 An. moucheti (W+)
DRC-MIK3 An. moucheti (W+)
KJ522813.1 Anopheles sp. 1 BSL-2014 (Kenya)
DRC-LWI3 An. sp. A/1 (W+)
DRC-LWI2 An. sp. A/1 (W+)
DRC-LWI1 An. sp. A/1 (W+)
DRC-KAT1 An. sp. A/1 (W+)
KR014819.1 Anopheles sp. NFL-2015 haplotype AN1 (Zambia)
DRC-KAT2 An. sp. A/1 (W+)
GUI-KAN1 An. sp. unknown (W-)
GUI-MAF1 An. sp. X (W+)
Anopheles theileri (3 sequences, Kenya / Zambia)
KJ522819.1 Anopheles sp. 7 BSL-2014 (Kenya)
KJ522818.1 Anopheles sp. 6 BSL-2014 (Kenya)
KJ522826.1 Anopheles sp. 14 BSL-2014 (Kenya)
KR014818.1 Anopheles cf. rivulorum NFL-2015 (Zambia)
JN994147.1 Anopheles cf. rivulorum LN-2012b (Zambia)
Anopheles rivulorum (3 sequences, S. Africa / Zambia)
MAD-AMB1 An. rufipes (W-)
MAD-AMB2 An. rufipes (W-)
JN994150.1 Anopheles rufipes (Zambia)
KR014828.1 Anopheles rufipes haplotype AN10 (Zambia)
KJ522822.1 Anopheles rufipes (Kenya)
Anopheles pretoriensis (3 sequences, Kenya / Zambia)
KR014835.1 Anopheles maculipalpis haplotype AN16 (Zambia)
JN994142.1 Anopheles maculipalplis (Zambia)
KJ522817.1 Anopheles maculipalpis (Kenya)
JN994143.1 Anopheles minimus C (Thailand)
Anopheles leesoni (3 sequences, Kenya / Zambia)
0.1

Figures

Figure 1. *Anopheles* species phylogenetic analysis of the *ITS2* gene. Maximum Likelihood molecular phylogenetic analysis of *Anopheles ITS2* sequences to demonstrate *An. species X* (red) phylogeny. The tree with the highest log likelihood (-3250.26) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 128 nucleotide sequences. There was a total of 169 positions in the final dataset. The *An. gambiae* complex is shown as a compressed sub-tree (blue). For comparison, various field-collected *Anopheles* sequences, previously obtained [24] are included, in addition to sequences obtained from GenBank, with their accession numbers provided.

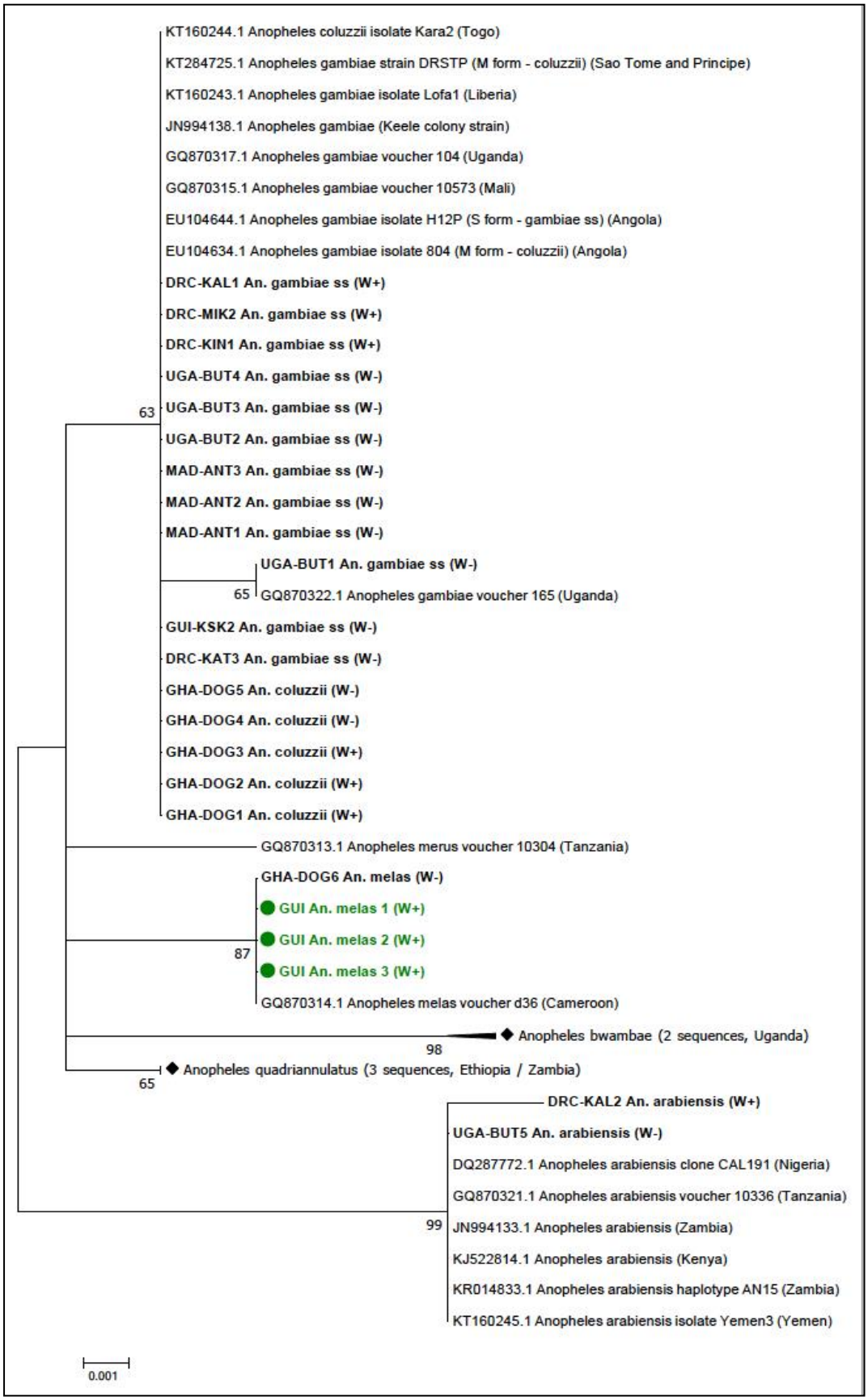

KT160244.1 Anopheles coluzzii isolate Kara2 (Togo)
KT284725.1 Anopheles gambiae strain DRSTP (M form - coluzzii) (Sao Tome and Principe)
KT160243.1 Anopheles gambiae isolate Lofa1 (Liberia)
JN994138.1 Anopheles gambiae (Keele colony strain)
GQ870317.1 Anopheles gambiae voucher 104 (Uganda)
GQ870315.1 Anopheles gambiae voucher 10573 (Mali)
EU104644.1 Anopheles gambiae isolate H12P (S form - gambiae ss) (Angola)
EU104634.1 Anopheles gambiae isolate 804 (M form - coluzzii) (Angola)
DRC-KAL1 An. gambiae ss (W+)
DRC-MIK2 An. gambiae ss (W+)
DRC-KIN1 An. gambiae ss (W+)
UGA-BUT4 An. gambiae ss (W-)
63
UGA-BUT3 An. gambiae ss (W-)
UGA-BUT2 An. gambiae ss (W-)
MAD-ANT3 An. gambiae ss (W-)
MAD-ANT2 An. gambiae ss (W-)
MAD-ANT1 An. gambiae ss (W-)
UGA-BUT1 An. gambiae ss (W-)
65
GQ870322.1 Anopheles gambiae voucher 165 (Uganda)
GUI-KSK2 An. gambiae ss (W-)
DRC-KAT3 An. gambiae ss (W-)
GHA-DOG5 An. coluzzii (W-)
GHA-DOG4 An. coluzzii (W-)
GHA-DOG3 An. coluzzii (W+)
GHA-DOG2 An. coluzzii (W+)
GHA-DOG1 An. coluzzii (W+)
GQ870313.1 Anopheles merus voucher 10304 (Tanzania)
GHA-DOG6 An. melas (W-)
GUI An. melas 1 (W+)
GUI An. melas 2 (W+)
87
GUI An. melas 3 (W+)
GQ870314.1 Anopheles melas voucher d36 (Cameroon)
Anopheles bwambae (2 sequences, Uganda)
98
Anopheles quadriannulatus (3 sequences, Ethiopia / Zambia)
65
DRC-KAL2 An. arabiensis (W+)
UGA-BUT5 An. arabiensis (W-)
DQ287772.1 Anopheles arabiensis clone CAL191 (Nigeria)
GQ870321.1 Anopheles arabiensis voucher 10336 (Tanzania)
99
JN994133.1 Anopheles arabiensis (Zambia)
KJ522814.1 Anopheles arabiensis (Kenya)
KR014833.1 Anopheles arabiensis haplotype AN15 (Zambia)
KT160245.1 Anopheles arabiensis isolate Yemen3 (Yemen)
0.001

Figure 2. *An. gambiae* complex species phylogenetic analysis of the *ITS2* gene. Maximum Likelihood molecular phylogenetic analysis of the *ITS2* gene for the *An. gambiae* complex, showing representative *w*AnMe-infected *An. melas* individuals from Guinea (green). The tree with the highest log likelihood (-785.62) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 45 nucleotide sequences. There was a total of 475 positions in the final dataset. For comparison, sequences from the *An. gambiae* complex, previously obtained [24], and from GenBank, with accompanying accession numbers, are included.

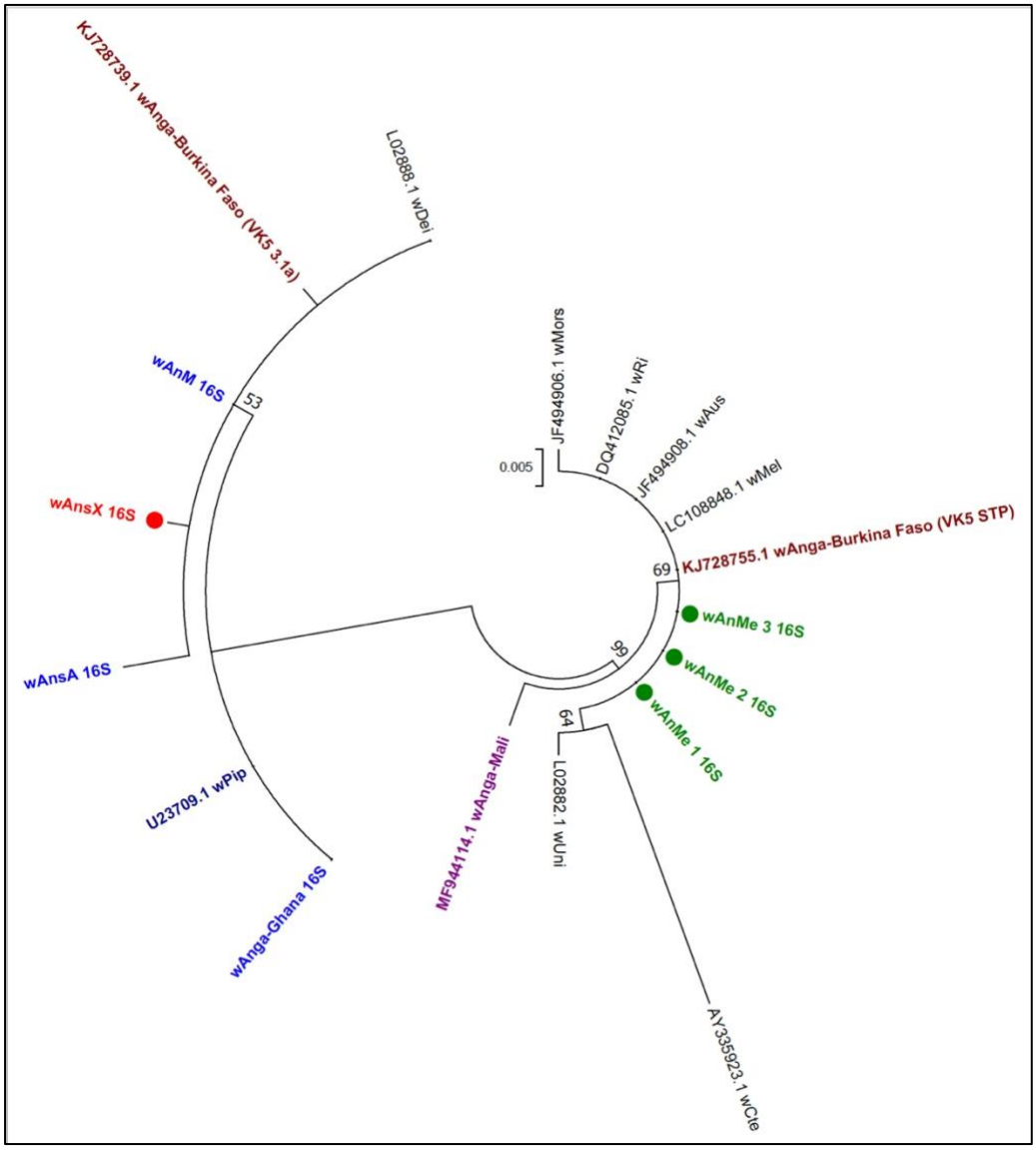


Figure 3. *Wolbachia* strain phylogenetic analysis using the *16S rRNA* gene. Maximum Likelihood molecular phylogenetic analysis of the *16S rRNA* gene for *w*AnsX-infected individual (red) and representative *w*AnMe-infected individuals (green) from Guinea. The tree with the highest log likelihood (-685.48) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 18 nucleotide sequences. There was a total of 338 positions in the final dataset. Sequences obtained from *Anopheles* species previously [24] are shown in blue. Accession numbers of additional sequences obtained from GenBank are shown, including *w*Pip (navy blue), *w*Anga-Mali (purple) and *w*Anga-Burkino Faso strains (maroon).

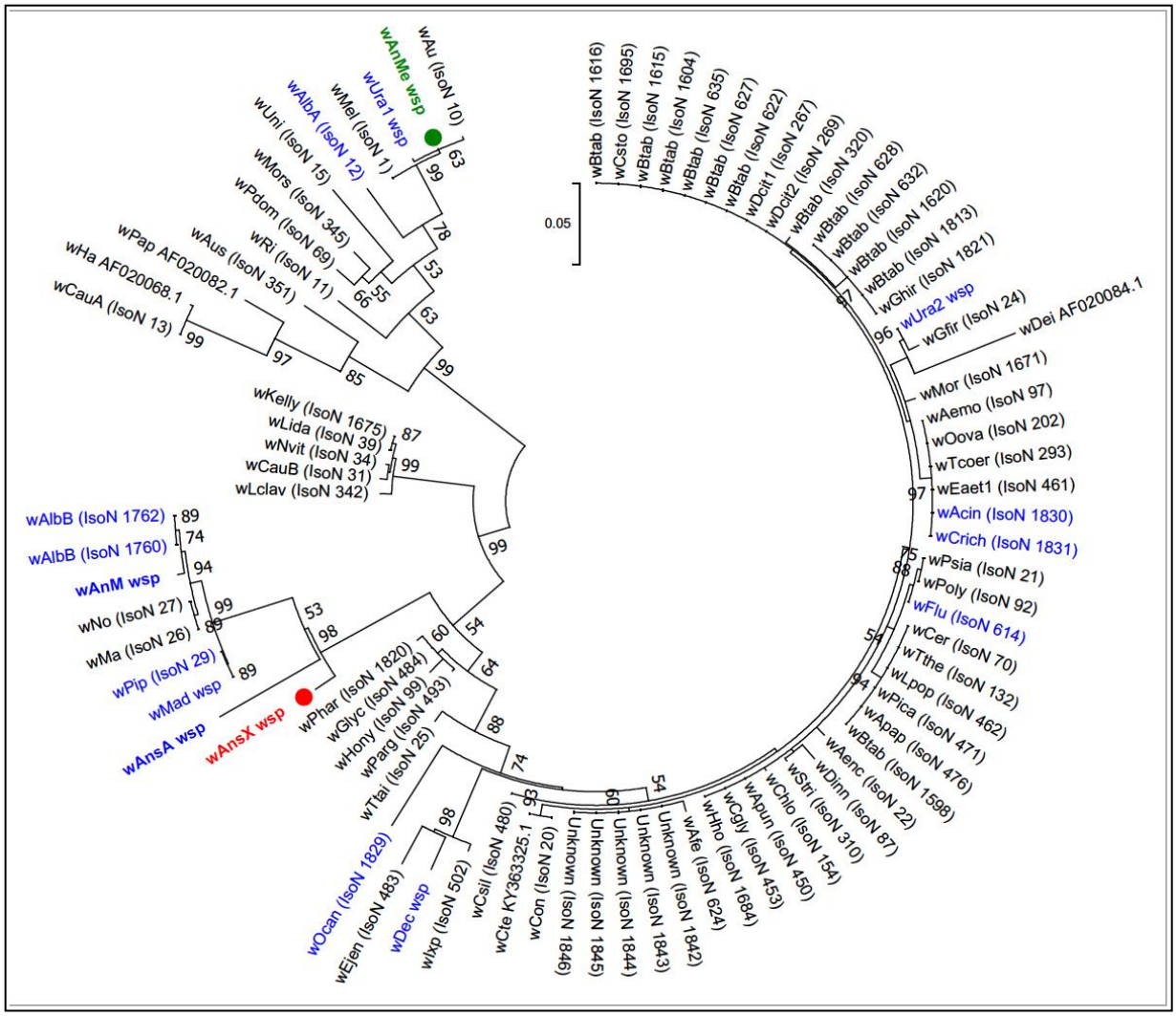


Figure 4. *Wolbachia* strain phylogenetic analysis using the *wsp* gene. Maximum Likelihood molecular phylogenetic analysis of the *wsp* gene for resident strains in *An. species X* (*w*AnsX, red) and *An. melas* (*w*AnMe, green) from Guinea. The tree with the highest log likelihood (-3646.57) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 86 nucleotide sequences. There was a total of 431 positions in the final dataset. Reference numbers of additional sequences obtained from the MLST database (IsoN; Isolate number) or GenBank (accession number) are shown. Strains isolated from mosquitoes are shown in blue, with those strains from other *Anopheles* species highlighted in bold.

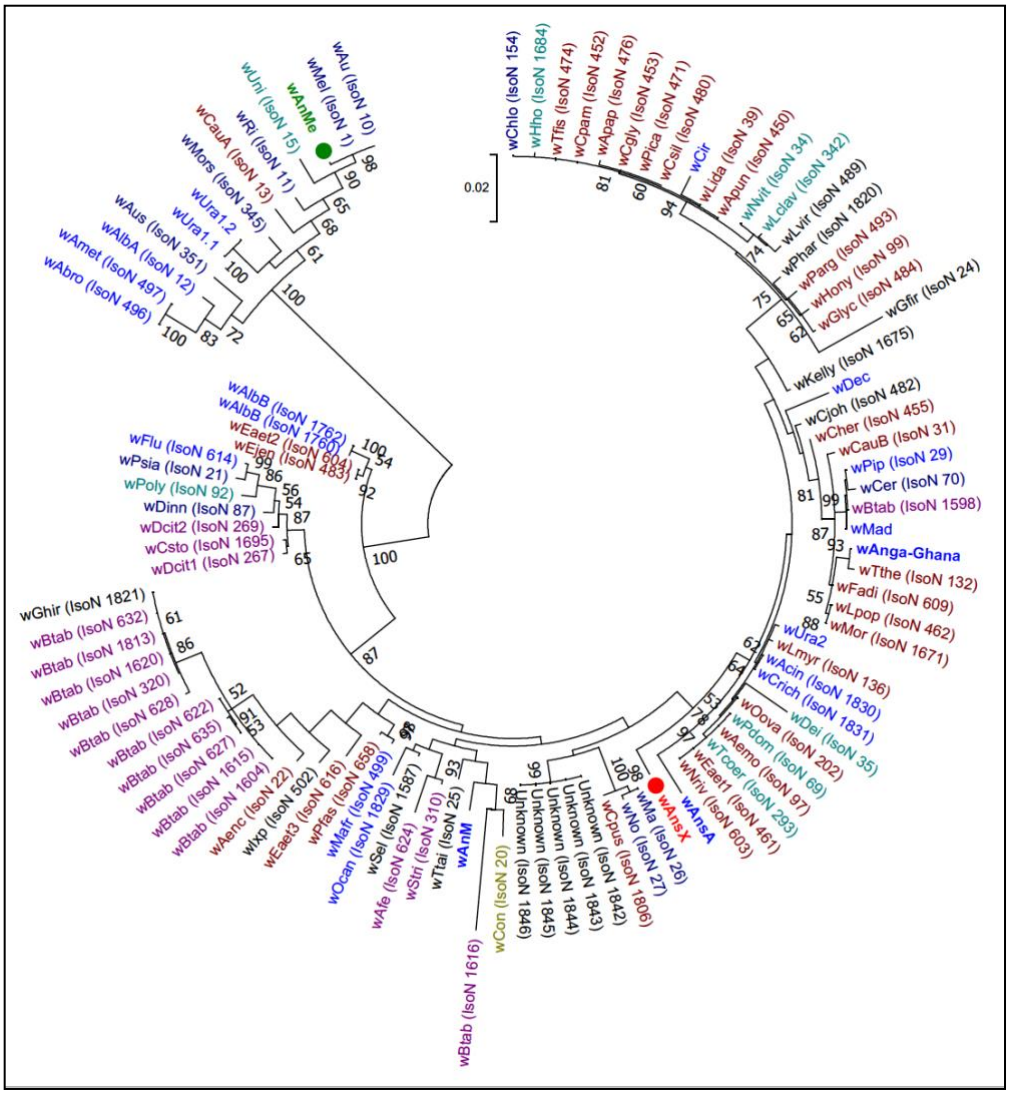


Figure 5. *Wolbachia* multilocus sequence typing (MLST) phylogenetic analysis of *Wolbachia* strains in *An. melas and An. species X.* Maximum Likelihood molecular phylogenetic analysis from concatenation of all five MLST gene loci for resident *Wolbachia* strains from *An. species X* (*w*AnsX; red) and *An. melas* (*w*AnMe; green). The tree with the highest log likelihood (-11404.41) is shown and drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 102 nucleotide sequences. There were a total of 2063 positions in the final dataset. Concatenated sequence data from *Wolbachia* strains downloaded from MLST database for comparison are shown with isolate numbers in brackets (IsoN). *Wolbachia* strains isolated from mosquito species are shown in blue, with those strains from other *Anopheles* species highlighted in 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.

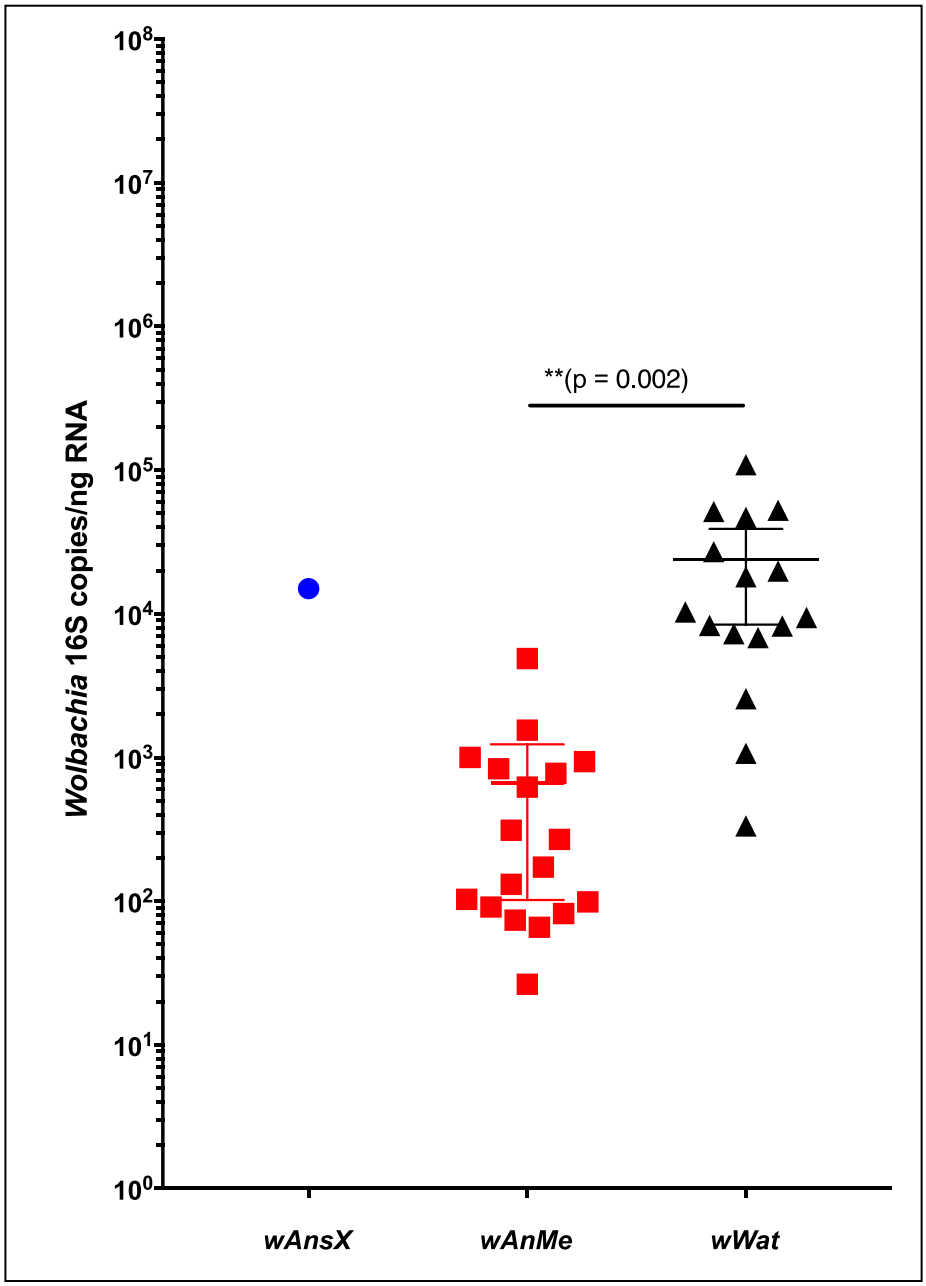


Figure 6. *w*AnsX, *w*AnMe and *w*Wat *Wolbachia* strain densities in wild caught female mosquitoes from the Maferinyah sub-prefecture of Guinea. Total RNA extracted from individual mosquitoes was standardised to 2.0 ng/µL prior to being used in qRT PCR assays targeting the *16S rDNA* gene. A synthetic oligonucleotide standard was designed to calculate *16S rDNA* gene copies per µL of RNA using a serial dilution series and all samples were run in triplicate in addition to no template controls.

Tables

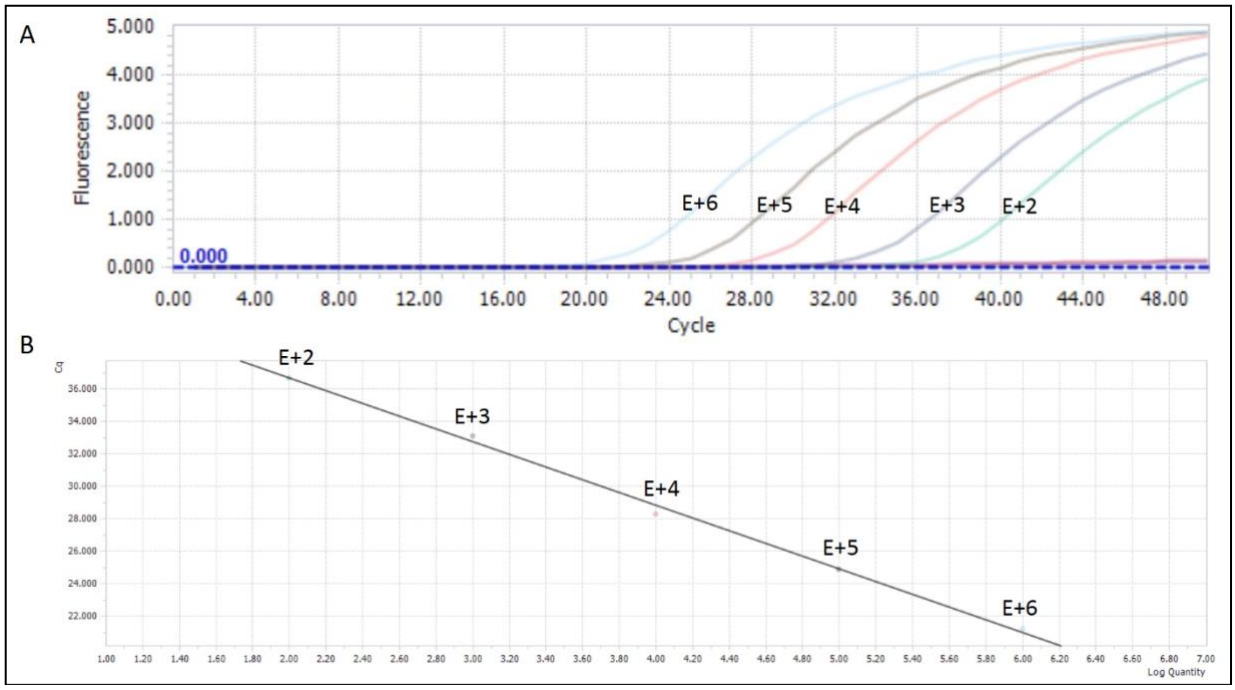
Table 1. *Wolbachia* prevalence rates in *Anopheles* species collected in two regions of Guinea in 2018. Species containing *Wolbachia*-infected individuals are denoted in bold.

Region	Collection location	Species	*Wolbachia*+/total individuals	Prevalence rate (%)
Faranah	Balayani	***An. gambiae* s.s.**	**4/139**	**2.9**
		An. coluzzii	0/1	0.0
		An. gambiae/coluzzii hybrid	0/1	0.0
		Unknown species	0/1	0.0
	Faranah	*An. gambiae* s.s.	0/26	0.0
		An. coluzzii	0/1	0.0
	Foulaya	*An. gambiae* s.s.	0/67	0.0
		An. gambiae/coluzzii hybrid	0/1	0.0
	Tindo	***An. gambiae* s.s.**	**1/48**	**2.1**
		An. coluzzii	0/1	0.0
		An. gambiae/coluzzii hybrid	0/2	0.0
Maferinyah	Fandie	*An. coluzzii*	0/22	0.0
		An. gambiae s.s.	0/2	0.0
		An. melas	0/3	0.0
	Maferinyah	*An. coluzzii*	0/7	0.0
		An. coustani	0/3	0.0
		An. gambiae s.s.	0/1	0.0
		An. squamosus	0/8	0.0
	Senguelen	*An. coluzzii*	0/26	0.0
		An. coustani	0/1	0.0
		An. gambiae s.s.	0/8	0.0
		An. melas	**18/168**	**10.7**
		An. species X	**1/1**	**100.0**
		An. squamosus	0/1	0.0
		An. gambiae/coluzzii hybrid	0/3	0.0

Table 2. Novel resident *Wolbachia* strain *WSP* typing and multilocus sequence typing (MLST) gene allelic profiles. Newly assigned novel alleles for *w*AnsX are shown in bold red font. **w*AnMe *hcpA* could not be assigned a novel allele number due to a possible double infection which was unresolvable, therefore the allele number of the closest match (CM) is shown with the number of single nucleotide differences to the closest match in brackets.

Mosquito species	***Wolbachia* strain**	***WSP* typing allele numbers**					**MLST gene allele numbers**				
		wsp	**HVR1**	**HVR2**	**HVR3**	**HVR4**	***gatB***	***coxA***	***hcpA***	***ftsZ***	***fbpA***
An. melas	*w*AnMe	23	1	12	21	19	1	1	CM1 (2)*	3	1
An. sp. X	*w*AnsX	**737**	**264**	**297**	3	**323**	**285**	**282**	**310**	**246**	**454**

Supplementary material



Supplementary Figure 1. *Wolbachia 16S rRNA* gene qRT PCR representative standard curve generated with a synthetic oligonucleotide standard through ten-fold serial dilutions (A) to measure 10E+6 to 10E+2 gene copies per µL (B).

Supplementary Table 1. **Additional sample details and ITS2 and CO1 GenBank accession numbers for mosquito species identification.** The species and target gene fragment are shown for *Wolbachia*-infected individuals and the accession number on GenBank.

Sample ID	Species	Gene fragment	Accession number
An. sp. X	*An. species X*	ITS2	XXXXX
An. melas1	*An. melas*	ITS2	XXXXX
An. melas2	*An. melas*	ITS2	XXXXX
An. melas3	*An. melas*	ITS2	XXXXX
Cx. watti1	*Cx. watti*	CO1	XXXXX

Supplementary Table 2. *Wolbachia* 16S GenBank accession numbers. Sample codes, *Wolbachia* strain names and GenBank accession numbers for Wolbachia 16S.

Sample ID	**Strain**	**16S**
An. melas1	*w*AnMe	XXXXX
An. melas2	*w*AnMe	XXXXX
An. melas3	*w*AnMe	XXXXX
An. sp. X	*w*AnsX	XXXXX

Supplementary Table 3. *Wolbachia wsp* and MLST gene GenBank accession numbers. *Wolbachia* strain names *Wolbachia* wsp, and MLST gene sequence Genbank accession numbers.

Strain	**wsp**	**gatB**	**coxA**	**hcpA**	**ftsZ**	**fbpA**
*w*AnMe	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
*w*AnsX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
*w*Wat	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX

Supplementary Table 4. qRT PCR data for *Wolbachia*-infected individuals. RNA extracts were added to Qubit™ RNA High Sensitivity Assays (Invitrogen) and total RNA was measured using a Qubit Fluorometer (Invitrogen). All RNA extracts were then diluted to produce extracts that were 2.0 nanograms (ng)/ µL prior to being used in qRT PCR) assays targeting the *16S rDNA* gene. A synthetic oligonucleotide standard (Integrated DNA Technologies) was designed to calculate *16S rDNA* gene copies per µL using a serial dilution series.

sample code	Insect species	Total RNA		*Wolbachia 16S* RT qPCR gene amplification										Asaia 16S RT qPCR gene amplification (Ct)
		ng/mL	ng/µL	replicate 1		replicate 2		replicate 3		mean Ct	standard deviation Ct	mean copies/µL	standard deviation copies/µL	
				Ct	copies/ul	Ct	copies/ul	Ct	copies/ul					
P2.AUG.E2	*An. species X*	3.40E+04	34.0	25.90	1.97E+04	26.43	1.43E+04	28.65	1.10E+04	26.99	1.46	1.50E+04	4.37E+03	34.92
P2.AUG.A1	*An. melas*	8.12E+04	81.2	34.09	1.40E+02	33.39	2.13E+02	35.81	1.65E+02	34.43	1.25	1.73E+02	3.71E+01	32.79
P2.AUG.B1	*An. melas*	8.10E+04	81.0	30.63	1.13E+03	30.69	1.09E+03	31.22	2.44E+03	30.85	0.32	1.55E+03	7.70E+02	32.61
P2.AUG.C1	*An. melas*	2.22E+04	22.2	31.47	6.80E+02	31.39	7.14E+02	31.93	1.61E+03	31.60	0.29	1.00E+03	5.28E+02	29.03
P2.AUG.B2	*An. melas*	4.34E+04	43.4	31.31	7.49E+02	30.75	1.05E+03	32.72	1.01E+03	31.59	1.02	9.38E+02	1.64E+02	30.74
P2.AUG.C2	*An. melas*	1.09E+04	10.9	31.24	7.82E+02	31.40	7.10E+02	32.71	1.02E+03	31.78	0.81	8.37E+02	1.62E+02	28.51
P2.AUG.D2	*An. melas*	1.28E+04	12.8	35.12	7.48E+01	34.69	9.71E+01	36.29	1.25E+02	35.37	0.83	9.89E+01	2.50E+01	32.72
P2.AUG.B9	*An. melas*	4.36E+04	43.6	28.00	5.54E+03	29.00	3.03E+03	29.66	6.10E+03	28.89	0.84	4.89E+03	1.64E+03	31.09
P2.AUG.B10	*An. melas*	2.38E+04	23.8	36.76	2.78E+01	36.82	2.68E+01	39.05	2.47E+01	37.54	1.31	2.64E+01	1.57E+00	27.50
P2.AUG.C10	*An. melas*	1.36E+04	13.6	32.95	2.78E+02	32.90	2.87E+02	34.44	3.70E+02	33.43	0.88	3.12E+02	5.04E+01	26.08
P2.AUG.E10	*An. melas*	1.49E+04	14.9	35.25	6.92E+01	35.47	6.06E+01	36.81	9.20E+01	35.84	0.84	7.39E+01	1.62E+01	26.96
P2.AUG.G10	*An. melas*	1.90E+04	19.0	35.23	7.00E+01	34.43	1.14E+02	36.30	1.24E+02	35.32	0.94	1.03E+02	2.88E+01	29.23
P2.AUG.H10	*An. melas*	1.51E+04	15.1	36.60	1.04E+02	35.30	6.71E+01	37.16	7.49E+01	36.35	0.95	8.20E+01	1.95E+01	28.55
P2.AUG.G11	*An. melas*	8.26E+04	82.6	34.70	9.65E+01	34.52	1.08E+02	35.60	1.87E+02	34.94	0.58	1.31E+02	4.93E+01	33.31
P2.AUG.H11	*An. melas*	3.12E+04	31.2	34.71	9.59E+01	36.42	3.41E+01	36.05	1.44E+02	35.73	0.90	9.12E+01	5.49E+01	31.12
P2.AUG.A12	*An. melas*	6.24E+04	62.4	31.79	5.61E+02	31.65	6.10E+02	32.49	1.16E+03	31.98	0.45	7.77E+02	3.33E+02	29.74
P2.AUG.B12	*An. melas*	6.40E+04	64.0	34.16	1.34E+02	33.24	2.33E+02	34.14	4.41E+02	33.85	0.53	2.69E+02	1.56E+02	31.72
P2.AUG.C12	*An. melas*	5.22E+04	52.2	33.83	5.29E+02	33.21	7.60E+02	33.67	5.81E+02	33.57	0.32	6.23E+02	1.22E+02	30.53
P2.AUG.G12	*An. melas*	4.38E+04	43.8	36.17	1.34E+02	38.37	3.68E+01	38.93	2.65E+01	37.82	1.46	6.58E+01	5.92E+01	21.41
***An. melas* (overall)**										**33.94**	**0.81**	**6.69E+02**	**2.34E+02**	**29.65**
CxWat1	*Cx watti*	9.24E+04	92.4	35.75	4.69E+02	38.38	1.21E+02	36.00	4.12E+02	36.71	1.45	3.34E+02	1.87E+02	
CxWat2	*Cx watti*	5.76E+04	57.6	36.04	4.03E+02	33.92	1.20E+03	33.38	1.59E+03	34.45	1.41	1.07E+03	6.06E+02	
CxWat3	*Cx watti*	1.46E+04	14.6	28.05	2.49E+04	27.53	3.25E+04	28.13	2.39E+04	27.90	0.33	2.71E+04	4.74E+03	
CxWat4	*Cx watti*	8.36E+04	83.6	29.98	9.19E+03	30.20	8.20E+03	30.43	7.29E+03	30.20	0.23	8.23E+03	9.52E+02	
CxWat5	*Cx watti*	1.46E+04	14.6	25.45	9.51E+04	25.93	7.42E+04	24.46	1.58E+05	25.28	0.75	1.09E+05	4.38E+04	
CxWat6	*Cx watti*	5.10E+04	51.0	25.85	2.14E+04	25.57	2.54E+04	26.69	1.28E+04	26.04	0.58	1.99E+04	6.45E+03	
CxWat7	*Cx watti*	6.90E+04	69.0	26.9	1.12E+04	27.09	1.00E+04	27.13	9.76E+03	27.04	0.12	1.03E+04	7.95E+02	
CxWat8	*Cx watti*	3.81E+04	38.1	26.7	1.27E+04	25.8	2.21E+04	25.99	1.96E+04	26.16	0.47	1.81E+04	4.86E+03	
CxWat9	*Cx watti*	6.80E+04	68.0	27.16	9.58E+03	28.11	5.35E+03	27.09	1.00E+04	27.45	0.57	8.31E+03	2.57E+03	
CxWat10	*Cx watti*	3.86E+04	38.6	29.17	2.79E+03	27.02	1.04E+04	27.34	8.58E+03	27.84	1.16	7.27E+03	3.99E+03	
CxWat11	*Cx watti*	4.29E+04	42.9	29.24	2.68E+03	28.92	3.26E+03	29.9	1.79E+03	29.35	0.50	2.57E+03	7.41E+02	
CxWat12	*Cx watti*	2.48E+04	24.8	24.74	4.23E+04	24.13	6.14E+04	24.43	5.11E+04	24.43	0.31	5.16E+04	9.59E+03	
CxWat13	*Cx watti*	5.60E+04	56.0	28.9	3.30E+03	27.39	8.32E+03	27.28	8.90E+03	27.86	0.91	6.84E+03	3.08E+03	
CxWat14	*Cx watti*	2.66E+04	26.6	24.67	4.41E+04	24.52	4.84E+04	24.51	4.87E+04	24.57	0.09	4.70E+04	2.54E+03	
CxWat15	*Cx watti*	5.70E+04	57.0	26.9	1.12E+04	27.58	7.41E+03	27.15	9.64E+03	27.21	0.34	9.43E+03	1.93E+03	
CxWat16	*Cx watti*	1.42E+04	14.2	24.25	5.71E+04	24.74	4.23E+04	24.21	5.85E+04	24.40	0.30	5.26E+04	8.99E+03	
***Cx. watti* (overall)**										**27.93**	**0.59**	**2.37E+04**	**5.99E+03**	