wMel Wolbachia genome remains stable after 7 years in Australian Aedes aegypti field populations

Kimberley R. Dainty1,2, Jane Hawkey3, Louise M. Judd2, Etiene C. Pacidônio1,4, Johanna M. Duyvestyn1,4, Daniela S. Gonçalves1,4, Silk Yu Lin1,4, Tanya B. O’Donnell1,4, Scott L. O’Neill1,4, Cameron P. Simmons1,4,6, Kathryn E. Holt3,6 and Heather A. Flores1,4,*

Abstract
Infection of wMel Wolbachia in Aedes aegypti imparts two signature features that enable its application for biocontrol of dengue. First, the susceptibility of mosquitoes to viruses such as dengue and Zika is reduced. Second, a reproductive manipulation is caused that enables wMel introgression into wild-type mosquito populations. The long-term success of this method relies, in part, on evolution of the wMel genome not compromising the critical features that make it an attractive biocontrol tool. This study compared the wMel Wolbachia genome at the time of initial releases and 1–7 years post-release in Cairns, Australia. Our results show the wMel genome remains highly conserved up to 7 years post-release in gene sequence, content, synteny and structure. This work suggests the wMel genome is stable in its new mosquito host and, therefore, provides reassurance on the potential for wMel to deliver long-term public-health impacts.

DATA SUMMARY
Raw sequence data and assembled genomes are available from the National Center for Biotechnology Information (NCBI) under BioProject accession number PRJNA695307.

INTRODUCTION
Dengue is the fastest growing mosquito-borne disease, having increased in incidence by 30-fold over the past 50 years [1]. Around 400 million people a year are estimated to be infected by dengue viruses, with over half of the world’s population at risk of the disease [2]. As historical methods to control arboviral disease transmission such as chemical insecticide-based programmes and breeding-site reduction struggle to compete with increasing urbanization, population density and global transportation systems, novel vector control methodologies are being developed to address this growing problem [3, 4].

A group of novel vector control methodologies focus on the use of the bacteria, Wolbachia pipientis. Wolbachia is a maternally inherited endosymbiont present in 40–70% of insects worldwide [5–7]. Many Wolbachia strains induce a reproductive manipulation called cytoplasmic incompatibility (CI). This provides a fitness advantage to Wolbachia-infected females, helping to drive Wolbachia into wild-type populations [8, 9]. More recently, multiple Wolbachia strains have been shown to provide host protection from pathogenic viruses [10–12]. The World Mosquito Program (WMP) transinfected a strain of Wolbachia native to Drosophila melanogaster, wMel, into Aedes aegypti. This infection causes the two desired features, CI [13, 14] and inhibition of arbovirus transmission [13, 15–22], which underpin the biocontrol method.

wMel-infected Ae. aegypti were first released in Australia in the Cairns suburbs of Gordonvale and Yorkeys Knob in 2011. After 10 weeks of releases, wMel was successfully introgressed into the wild-type population and has remained at high
Impact Statement

*Aedes aegypti* mosquitoes transmit a number of arboviruses that cause human disease, including dengue. The introduction of the wMel strain of *Wolbachia* into *A. aegypti* populations has proven to be an effective biocontrol method for dengue in Cairns, Australia, where it was first established in 2011. The infection of wMel into *A. aegypti* significantly reduces the ability of the mosquito to transmit virus between humans. It also causes a reproductive manipulation that enables introduction of wMel into wild-type mosquito populations. These two features must remain stable in field conditions for the continued success of this intervention. Here, we examine the genomic evolution of wMel in *A. aegypti* since its establishment in Cairns, Australia. By using two sequencing methods, we are able to examine the gene sequence, content and structure. We find that the wMel genome has remained highly conserved in *A. aegypti*, since its establishment in Cairns, Australia. This work gives reassurance on the long-term applicability of wMel as a biocontrol method for arboviruses such as dengue.

**METHODS**

**Sample collection and processing**

Samples from 2011 were acquired via ovitraps deployed March–April 2011. *Ae. aegypti* eggs were collected from Gordonvale, Cairns. Eggs were hatched at 26 °C, 60% relative humidity with a 12 h light:dark cycle, and reared to second instar larvae before being stored in 80% ethanol at 4 °C. A total of 10 larval samples were sequenced (Table S1, available with the online version of this article).

Dainty et al., Microbial Genomics 2021;7:000641
Samples from 2018 were acquired via ovitraps deployed March–May 2018. *Ae. aegypti* eggs were collected from four release suburbs in Cairns (Gordonvale, Yorkeys Knob, Mount Sheridan, Smithfield). Eggs from each ovitrap were hatched and reared at 26 °C, 60% relative humidity with a 12h light-dark cycle. Mosquitoes were aged at least 7 days post-emergence when their ovaries were obtained to enrich for *Wolbachia*, as wMel density is substantially higher in this tissue [10]. Ovaries from 2 to 30 *Ae. aegypti* females per trap were dissected and frozen in liquid nitrogen before being stored at ~80 °C. As ovitraps are likely to contain full-sibling individuals [40, 41], each of these collections are considered to represent the offspring from a single *Wolbachia*-infected mother. A total of 30 samples across the four suburbs were sequenced (Table S1).

Samples were homogenized in buffer ATL using a hand-pestle, before DNA was extracted using the MagAttract HMW DNA kit (Qiagen) with a 50 µl elution in nuclease-free water, following the manufacturer’s instructions. Libraries were prepared for each sample with the Nextera DNAFlex library prep kit (Illumina) using unique index tags. Nextera DNAFlex libraries were prepared according to the manufacturer’s directions with one significant deviation, all reactions were scaled to 25% of recommended volumes. This change did not significantly affect the performance of the library preparation. Samples were sequenced via Illumina platforms, generating 150-base-paired-end reads. Details regarding samples and associated sequencing data are available in Table S2.

For long-read sequencing, library preparation was performed using the ligation-based kit (LSK109) with native barcoding (NBD103) to multiplex samples (Oxford Nanopore Technologies), before being loaded onto a R9.4 flow cell (Oxford Nanopore Technologies) and sequenced using the MinION device (Oxford Nanopore Technologies), as described previously [42]. DNA was not sheared or size selected prior to library preparation. Reads were basecalled with Guppy v3.6.0 using the R9.4.1 450bps HAC (high accuracy) model. Reads were filtered using Filllong (https://github.com/rrwick/Filtlong) using default parameters to keep reads at least 1000bp long before further analysis. Details regarding samples and associated sequencing data are available in Table S2.

**Analysis**

BioBloom Tools v2.3.2 [43] was used to identify Illumina reads belonging to *Ae. aegypti* (accession no. NC_035107.1) using default parameters, and these reads were excluded from downstream analysis. Single nucleotide variants were identified via mapping the remaining Illumina reads to the *D. melanogaster* wMel reference genome [28] (NC_002978.6) using the RedDog v1b.11 pipeline (https://github.com/katholt/reddog) according to the developers’ guidelines and using standard parameters. Reads were not trimmed prior to analysis. Briefly, Bowtie2 version 2.2.3 [44] was used to map reads to the reference sequence, before SAMtools version 1.9 [45] called SNPs with QUAL values ≥30. For 2011 samples, Illumina read depth varied between 16.58× and 97.78× (mean 41.84×). Coverage was 100% for 9/10 genomes, and 99.99% for the remaining sample. For 2018 samples, depth ranged between 73.17× and 448.18× (mean 179.69×), with coverage of 100% achieved for all genomes. Visualization of the polymorphisms across the genome was created using CiVi [46]. Low-frequency variants were detected in samples from 2018 with LoFreq v2.1.3.1 [47] using standard parameters. Variants identified with a strand bias above 10 were removed from the final data set. Gene copy number variation was assessed by normalizing the sequencing depth of coverage for each gene in the sequenced genomes. Normalization was calculated by dividing the mean depth of coverage for each gene by the mean depth of coverage for the whole genome.

Insertion sequence (IS) elements were identified by searching against the ISfinder [48] database via the Issaga [49] web server (available at http://issaga.biotoul.fr/issaga_index.php) using the wMel reference genome. Identified IS queries that had greater than 80% similarity (Table S3) were mapped to the short-read sequencing data for each sample using ISMapper [50] with default parameters to identify IS insertion sites. Insertion or deletion of IS elements identified as imprecise (*) or uncertain (?) were considered mapping artefacts.
A hybrid assembly was performed for all samples using all long-reads, and non-\textit{Ae. aegypti} Illumina reads using Unicycler V0.4.7 [51] with default parameters. Long-reads were not pre-filtered to exclude \textit{Ae. aegypti} reads, due to the lower accuracy of these reads being unsuitable for use with BioBloom Tools. However, Unicycler works to first assemble short-reads (pre-filtered) before scaffolding these assemblies together with the long-reads, meaning long-reads belonging to \textit{Ae. aegypti} are essentially ignored. Genomes were trimmed and rotated to the \textit{dnaA} gene, then confirmed closed in Unicycler. Single circularized genomes were assembled for seven samples from 2018. Gene synteny was assessed using Progressive Mauve software [52] (available at http://darlinglab.org/mauve/mauve.html), using standard parameters.

**RESULTS**

The first releases of wMel-infected \textit{Ae. aegypti} began in January 2011 in the Cairns, Australia, suburbs of Gordonvale and Yorkeys Knob (Fig. 1). In order to identify changes in the wMel genome since its transinfection from \textit{D. melanogaster} to \textit{Ae. aegypti}, wMel-infected larval samples collected from Gordonvale, Cairns, in 2011 were sequenced (Fig. 2). Analysis of wMel \textit{Wolbachia} genomes collected from Gordonvale ascertained the presence of two SNPs, compared to the previously described reference genome of wMel (sequenced directly from its native host, \textit{D. melanogaster}) (Fig. 3, Table 1). The first SNP, labelled SNP_A, was present in all 10 samples, and represents an intergenic T-A change at position 1097797.

**Fig. 2.** Timeline of the establishment and sampling of wMel described in this study. The \textit{Wolbachia} strain wMel was purified from \textit{D. melanogaster} and transinfected into the \textit{Ae. albopictus}-derived cell line RML12. After approximately 24 months of serial passaging, the wMel was transinfected into \textit{Ae. aegypti} mosquitoes. Releases of wMel-infected mosquitoes into the suburbs of Gordonvale (GV) and Yorkeys Knob (YK), and Mount Sheridan (MS) and Smithfield (SF), began approximately 42 and 115 months post-initial transfection, respectively. We sequenced the genomes of wMel from mosquitoes at two timepoints, 2011 and 2018, the numbers of genomes sequenced from each suburb are indicated under the suburb initials and year.

**Fig. 3.** wMel \textit{Wolbachia} genome stability over time. (a) Progenitor wMel strain sequenced from \textit{D. melanogaster} in 2004. (b) wMel strain in released \textit{Ae. aegypti} sampled in 2011. (c) wMel strain in released \textit{Ae. aegypti} sampled in 2018. Circles correspond to the following: blue layer, forward strand genes; green layer, reverse strand genes; orange segments, phage regions – in clockwise order WO-A, pyocin-like, WO-B, WO-B. SNPs are designated letter labels (A–F) with mutation outcome listed as either intergenic, synonymous (SYN) or non-synonymous (NON-SYN), as well as number of samples in which the SNP was found. *SNP_A is an error in the reference sequence rather than a novel SNP since wMel transfer to \textit{Ae. aegypti}.
of the wMel reference genome. The second SNP found in the 2011 samples, SNP_B, was present in 3 of 10 samples, and represents a synonymous T-C change in the hypothetical protein WD1228 at position 1174712. Seven indel events were also identified compared to the reference genome, each present in all 10 samples (Table 2). Short-read sequencing data was also used to assess IS movement within the genomes using the ISsaga and ISMapper programmes. No evidence of novel insertions or deletions of any IS elements were identified. Finally, gene copy number variation was assessed by normalizing the mean sequencing depth for each gene by the mean depth coverage of the entire genome. No distinct change in gene copy number was observed in any of the sequenced genomes (Fig. S1a). These data suggest low levels of wMel genome polymorphism have occurred as a result of host transfer.

Since 2011, releases of wMel-infected Ae. aegypti populations have occurred throughout the regional cities of Cairns, including Mount Sheridan and Smithfield, which were part of releases completed in 2017 (Figs 1 and 2). Longitudinal monitoring of Ae. aegypti wMel infection frequency indicates Wolbachia introgression has remained stable throughout these areas since establishment [26]. wMel-infected mosquitoes were collected in 2018 from Gordonvale and Yorkeys Knob in order to ascertain genomic changes accrued in the field, in the 86 months since the end of releases. Alongside this, wMel-infected mosquitoes from Mount Sheridan and Smithfield were also collected in 2018. These samples represented a shorter time in the field of 13 and 10 months, respectively, but may have been more diverse by virtue of genetic drift due to the wMel mosquito line having been maintained as a smaller laboratory population until release (Fig. 2).

All samples from 2018 were sequenced and analysed in the same manner as those from 2011. Again, SNP_A was identified in all samples from 2018 (Fig. 3, Table 1). Interestingly, SNP_B was observed in 9/30 samples, a similar proportion as in 2011, and was located across all four suburbs (13–50%, depending on suburb). A further four SNPs were identified across the 2018 samples, but were only present in one or two samples. SNP_C, E and F were each identified in one sample from Smithfield, Yorkeys Knob and Mount Sheridan, respectively, and cause non-synonymous, intergenic and synonymous changes. SNP_D, which causes a non-synonymous change in hypothetical protein WD_0244, was observed in two samples, one from Mount Sheridan and one from Yorkeys Knob. As WD_0244 represents a small hypothetical protein, one concern was that it has been misannotated. However, a study looking at wMel gene expression across D. melanogaster developmental stages showed WD_0244 to be expressed across multiple life stages, supporting its current annotation [53]. Of these novel four SNPs, no greater proportion was observed in the older sites, Gordonvale and Yorkeys Knob, compared to the more recent sites, Smithfield and Mount Sheridan (Table S4). Each of the seven indel differences identified in the 2011 samples were again present in all of the 2018 samples (Table 2). As with 2011 samples, no evidence of novel insertions or deletions of any IS elements were

---

**Table 1. SNPs present in wMel Wolbachia from Ae. aegypti collected in 2011 and 2018 compared to the wMel reference sequence from D. melanogaster.**

| SNP   | Position | Reference nucleotide | Alternate nucleotide | Change outcome | Gene accession | Ancestral codon | Derived codon | Ancestral amino acid | Derived amino acid | Gene product |
|-------|----------|----------------------|---------------------|----------------|----------------|----------------|--------------|---------------------|------------------|-------------|
| SNP_A | 097797   | T                    | A                   | Intergenic     | -              | -              | -            | -                   | -                | -           |
| SNP_B | 1174712  | C                    | T                   | Synonymous     | WD_1228        | GGA            | CAT          | R                   | H                | Hypothetical protein |
| SNP_C | 17401    | A                    | G                   | Non-synonymous | WD_0017        | AGA            | CAT          | Q                   | S                | Hypothetical protein |
| SNP_D | 229585   | C                    | A                   | Non-synonymous | WD_0244        | CAG            | TCA          | Q                   | S                | Hypothetical protein |
| SNP_E | 56862    | A                    | G                   | Intergenic     | WD_0790        | -              | -            | -                   | -                | -           |
| SNP_F | 735666   | G                    | A                   | Synonymous     | WD_0244        | TCG            | TCG          | S                   | S                | DNA polymerase III, α subunit |

SNP_A is an error in the reference sequence rather than a novel SNP since wMel transfer to Ae. aegypti.
identified. Again, no distinct change in gene copy number variation was observed in any of the sequenced genomes (Fig. S1b–e). These data suggest field establishment is having little effect on the wMel genome.

While the goal of our study was to identify SNPs that could have a potential impact on wMel-induced phenotypes, we also identified low-frequency variants within individual samples as this would provide an indication of the mutational input on which selection and drift can act. LoFreq was used to identify low-frequency SNPs from the short-read sequencing data for all 2018 samples (File S1). We identified a total of 281 variants across all samples in addition to our previously identified SNPs. Of the variants identified, 74% had an allele frequency <10%, and 98.6% had an allele frequency <20%. Frequency variation was also assessed for the six fixed SNPs identified in this study across all genomes sequenced (Fig. S2). SNP_B, which was found to be fixed in 12 of 40 genomes sequenced, was found at lower frequencies in a further nine genomes from the 2018 collections. SNP_E and SNP_F, which were each found to be fixed in 1 of 40 sequenced genomes, were found at lower frequencies in one additional genome each from the 2018 collections. As all samples from 2018 were generated by pooling material, this represents the frequency of the pool, rather than heteroplasmy within individual mosquitoes.

Long-read sequencing was undertaken on the 2018 samples collected from Gordonvale using the Oxford Nanopore platform to allow for the identification of large structural rearrangements. The Unicycler program was used to perform hybrid de novo genome assemblies using both the short- and long-read data. Seven 2018 Gordonvale genomes were able to be completely resolved, and genomes closed. The Mauve program was used to identify changes to gene order or genome rearrangement in the seven resolved genomes from Gordonvale in comparison to the D. melanogaster wMel reference genome. No changes were observed in any of the genomes compared to the reference genome (Fig. S3). Each genome aligned in full to every other genome, creating single homologous blocks that indicate the genomes are internally free from genomic rearrangement.

DISCUSSION

The wMel strain of Wolbachia has been introgressed into field populations of Ae. aegypti since 2011 as an intervention for reducing the prevalence of arboviruses. Observational data from across Australia show a reduction in dengue incidence since the introgression of Wolbachia into Cairns and Townsville [26, 27]. This intervention has proven effective due to the maintenance of viral inhibition and CI caused by wMel infection after its artificial transfer from its native host, D. melanogaster, to Ae. aegypti [13]. However, previous transfections using a closely related Wolbachia strain, wMelPop, resulted in substantial changes to its genome after host transfer. Therefore, we hypothesized that this host transfer, along with the introgression of wMel into wild Ae. aegypti populations with much larger population sizes than lab-reared colonies, may have caused substantial genomic changes. Our results, however, have shown a remarkably stable genome.

A total of six SNPs were identified across the 2011 and 2018 field samples. Only one SNP (SNP_A, T → A) was present in all 2011 and 2018 samples. However, when queried further, this position was found to contain an ‘A’ in 100% of genome sequences in three separate studies of wMel genomes from D. melanogaster [38, 39, 54]. Therefore, we obtained the Sanger sequencing trace data from the original wMel reference genome. The Sanger traces indicated a mismatch at this position, with trace data indicating a clear ‘A’ base, rather than the called ‘T’ base (Fig. S4a). In combination, this data provides clear evidence that SNP_A is instead an error in the reference genome, rather than a substitution that has occurred since transinfection of wMel into Ae. aegypti. The seven indel events identified in this study were also present in all 2011 and 2018 sequences. Each of these was again examined in the Sanger sequencing trace data from the original wMel reference genome. For six out of seven of these indel events (positions 1006081, 1020475, 1094458, 1161850, 1163170, and 1177853) there is strong evidence they are errors in the reference sequence (Fig. S4b–d, f–h). For the other indel at position 1103468, sequencing traces do not clearly indicate the number of Ts present at the position (Fig. S4e). However,

### Table 2. Indel differences identified between the wMel reference genome from D. melanogaster and the genomes sequenced in this study

| Position | Reference nucleotide | Alternate nucleotide | ORF effect | Gene | Gene product |
|----------|----------------------|----------------------|------------|------|--------------|
| 1006081  | GT                   | GTT                  | Frameshift | WD_1044 | Hypothetical protein |
| 1020475  | CTTTT               | CTTTTT              | Intergenic | –     | –            |
| 1094458  | GTT                 | GTTT                | Frameshift | WD_1143 | ispD/ispF bifunctional enzyme |
| 1103468  | ATTTT               | ATTT                | Frameshift | WD_1155 | Hypothetical protein |
| 1161850  | GTTTTT              | GTTTTTT            | Intergenic | –     | –            |
| 1163170  | GC                  | G                   | Frameshift | WD_1215 | Hypothetical protein |
| 1177853  | ACC                 | AC                  | Frameshift | WD_1231 | Hypothetical protein |

These indels are errors in the reference sequence rather than novel indels since wMel transfer to Ae. aegypti.
another published wMel genome sequence from D. melanogaster reports only three Ts at this position, rather than the four present in the reference genome, supporting this indel is also likely an error [55]. Given these results, it is clear these indel events are again errors in the reference sequence, rather than mutations that have occurred since transinfecion of wMel into Ae. aegypti.

The only other SNP identified in both 2011 and 2018 (SNP_B) represents a synonymous change in a hypothetical protein. As it was identified at a similar frequency at both time points, and present in all sampled suburbs, it is likely this change was present in the mosquito release colony and is providing no strong evolutionary benefit or disadvantage. The remaining four SNPs identified in this study were present in only 2018, and in one or two samples. The SNP found in two samples, SNP_D, was observed in two suburbs separated from one another by approximately 17 km (Yorkeys Knob and Mount Sheridan), making migration of mosquitoes between them unlikely. Unintentional human transportation of mosquitoes carrying this SNP may have occurred, however, or alternatively this SNP may have arisen independently, twice, potentially indicating this to be an adaptation to field inhabitancy. No evidence of novel insertions or deletions of any IS elements were identified in any of the sequences generated in this study. However, a higher depth of sequencing would increase possible avenues to explore this, as previous studies show the confidence in IS annotation increases with sequencing depth for the type of analysis performed in this study [50].

Our 2018 field sampling included mosquitoes from sites established in 2011, as well as 2017. Interestingly, the older release sites, Gordonvale and Yorkeys Knob, exhibited no greater abundance of genomic changes than the later release sites, Smithfield and Mount Sheridan, although the total number of polymorphisms is low. The two later release sites studied, Smithfield and Mount Sheridan, share no common polymorphisms despite these releases occurring around the same time and from the same colony stock. This suggests polymorphisms have not accumulated in the lab colony stock since the initial field releases in 2011. Importantly, no changes were identified in the genes known to be responsible for the CI phenotype, cifA and cifB [32–34, 56], indicating a stability of this desired trait. As the mechanism of Wolbachia-induced viral inhibition is yet unknown, no such conclusion can be made for this phenotype. However, the general stability of the genome predicts the trait is not at high risk of being lost due to changes in the Wolbachia genome. Consistent with this, a stable virus blocking phenotype was reported in field-derived Ae. aegypti, sampled from Cairns, Australia, 12 months post-release in 2012 [25].

Huang et al. [29] recently reported the sequencing of wMel from colony material, as well as wMel-infected mosquitoes collected 8 years post-release in Australia. While the general observations of wMel genome stability are observed in both studies, our sampling and sequencing methods have allowed for a more in-depth analysis of wMel genome evolution. Our baseline sampling comprised 10 individual larvae collected in 2011, whereas Huang et al. used a single pool of 400 mosquitoes from material that was collected from the field in 2013 and reared in the laboratory until its sequencing in 2019. This allowed us to identify the polymorphism labelled SNP_B in both 2011 as well 2018 samples, whereas Huang et al. 2020 only identified this SNP in 2019. Our use of 2011 material allowed us to confirm that SNP_B was present in wMel release material and did not originate in the field. Our study also found three additional SNPs in 2018 sampling (SNP_C, E and F) compared to that of Huang et al. study in 2019. However, only two of the suburbs assessed overlapped in the two studies. Thus, these SNPs may be unique to the suburbs sampled in this study as each were present at low frequency. Finally, our study also included the use of long-read sequencing, allowing us to fully resolve wMel genomes and conclude that no genome rearrangements have occurred since transfer to Ae. aegypti or release in the field.

This low level of variation coupled with the lack of IS movement and genomic rearrangements in the wMel genome is somewhat surprising given the occurrence of host transfer and introduction to a novel environment. A number of genomic changes were observed in the closely related strain of Wolbachia, wMelPop, after it underwent transinfection into a novel host. This strain, however, was transinfected into an additional Aedes albopictus-derived cell line, Aa23, which may have provided additional stresses leading to the genomic changes [30]. When low-frequency variation was assessed in this study, a large number of variants were identified; however, the majority of variants were at a low frequency. This suggests there is mutational input upon which selection can act; however, these are not yet fixing in the population. Similar findings to ours have been reported for the Wolbachia strain wCer2, which displayed stability of genome content and synteny, and low levels of sequence polymorphism in multiple novel hosts [57]. Alongside this, studies of the aphid endosymbiont Buchnera across different aphid species have reported low levels of DNA polymorphism, which was posited to be more likely shaped by symbiosis effects of aphids and Buchnera (bottleneck effects during maternal transmission and population fluctuations) rather than by features incidental to different aphid species [36, 58, 59].

The low level of variation observed here could be due to multiple explanations. It is possible that the novel host of wMel, Ae. aegypti, does not differ enough in terms of population dynamics and vertical endosymbiont transfer from its predecessor, D. melanogaster, to drive selective pressures on the wMel genome. Additionally, it is possible that the observed low-frequency mutations are under strong purifying selection, or that population bottlenecks associated with maternal transmission are limiting their inheritance. Furthermore, considering the relatively short time wMel had been established in the field when sequenced (∼68 months), coupled with the reportedly low estimated mutation rate of wMel (6.8×10^−10 substitutions per site per D. melanogaster generation [38, 39]), more time may be required to amass greater differences. Using the mutation rate of 6.8×10^−10 substitutions per site per generation and the estimated number of 104 generations

Dainty et al., Microbial Genomics 2021;7:000641
(assuming one generation per month) that wMel has been in *Ae. aegypti* (Fig. 2), we would expect to observe 0.0897 base substitutions across each genome. This estimation is consistent with what we have observed in this study. It is also known that the wMel genome has a number of partially intact DNA repair genes [28, 60], possibly limiting the number of polymorphisms that are able to fix in the germline.

The multi-faceted sequencing and analysis techniques used in this study have allowed for the establishment of complete wMel *Wolbachia* sequences and demonstrated a remarkably stable genome in terms of sequence, gene content, and structure. These results provide some of the first data regarding the genome stability of wMel. This, combined with recent field entomology data showing the stability of wMel-infection prevalence, gives reassurance on the potential for wMel to deliver long-term public-health impacts. Future studies should continue to monitor both the genomic evolution of the wMel genome, as well as the phenotypic features of viral inhibition and CI.

**Funding information**

This work was supported by Wellcome Trust Innovations Award no. 212914/Z/18/B (https://wellcome.org/), National Health and Medical Research Council (NHMRC) Award no. 1 132 412 (https://nhmrc.gov.au), and an Australian Government Research Training Program (RTF) scholarship (https://www.education.gov.au/research-training-program). The funders had no role in study design, data collection and analysis, decision to publish nor preparation of the manuscript.

**Acknowledgements**

We would like to thank Michael Butterworth, Robyn Gray, Mandi Soymonoff, Rod Bagita, Katya Salvador, Davide Bott, Elizabeth Surarez and Andrew Turley for technical assistance. We also thank two anonymous reviewers for valuable suggestions to improve the manuscript.

**Author contributions**

Conceptualization: C.P.S., K.E.H., H.A.F. Formal analysis: K.R.D., J.H. Funding acquisition: K.R.D., S.L.O.N. Investigation: K.R.D., L.M.J., E.C.P., J.M.D., D.S.G., T.B.O.D., S.Y.L. Methodology: K.R.D., J.H., K.E.H., H.A.F. Supervision: C.P.S., K.E.H., H.A.F. Writing – original draft: K.R.D. Writing – review & editing: K.R.D., C.P.S., K.E.H., H.A.F.

**Conflicts of interest**

The authors declare that there are no conflicts of interest.

**References**

1. Pang T, Mak TK, Gubler DJ. Prevention and control of dengue – the light at the end of the tunnel. *Lancet Infect Dis* 2017;17:e79–e87.
2. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, et al. The global distribution and burden of dengue. *Nature* 2013;496:504–507.
3. Mayer SV, Tesh RB, Vasilakis N. The emergence of arthropod-borne viral diseases: a global prospective on dengue, chikungunya and zika fevers. *Acta Trop* 2017;166:155–163.
4. Bonizzoni M, Gasperi G, Chen X, James AA. The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives. *Trends Parasitol* 2013;29:460–468.
5. Weinert LA, Araujo EV, Ahmed MZ, Welch JJ. The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc Biol Sci* 2015;282:20150249.
6. 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;7:e35844.
7. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. How many species are infected with *Wolbachia*? – a statistical analysis of current data. *FEMS Microbiol Lett* 2008;281:215–220.
8. Dobson SL, Fox CW, Jiggins FM. The effect of Wolbachia-induced cytoplasmic incompatibility on host population size in natural and manipulated systems. *Proc Biol Sci* 2002;269:437–445.
9. Serbus LR, Casper-Lindley C, Landmann F, Sullivan W. The genetics and cell biology of *Wolbachia*-host interactions. *Annu Rev Genet* 2008;42:683–707.
10. Fraser JE, De Bruyne JT, Iturbe-Ormaetxe I, Stepnell J, Burns RL, et al. Novel *Wolbachia*-transinfected *Aedes aegypti* mosquitoes possess diverse fitness and vector competence phenotypes. *PLoS Pathog* 2017;13:e1006751.
11. Hedges LM, Brownlie JC, O’Neill SL, Johnson KN. *Wolbachia* and virus protection in insects. *Science* 2008;322:702.
12. Teixeira F, Ferreira A, Ashburner M. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol* 2008;6:e1000002.
13. Walker T, Johnson PH, Moreira LA, Iturbe-Ormaetxe I, Freniti FD, et al. The wMel Wolbachia strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* 2011;476:450–453.
14. 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:e1006815.
15. Aliota MT, Walker EC, Yepes AU, Veliz ID, Christensen BM, et al. The wMel strain of *Wolbachia* reduces transmission of Chikungunya virus in *Aedes aegypti*. *PLoS Negl Trop Dis* 2016;10:e0004677.
16. Aliota MT, Peinado SA, Veliz ID, Osorio JE. The wMel strain of *Wolbachia* reduces transmission of Zika virus by *Aedes aegypti*. *Sci Rep* 2016;6:28792.
17. van den Hurk AF, Hall-Mendelin S, Pyke AT, Freniti FD, McElroy K, et al. Impact of *Wolbachia* on infection with chikungunya and yellow fever viruses in the mosquito vector *Aedes aegypti*. *PLoS Negl Trop Dis* 2012;6:e1892.
18. Carrington LB, Tran BCN, Le NTH, Luong TTH, Nguyen TT, et al. Field- and clinically derived estimates of *Wolbachia*-mediated blocking of dengue virus transmission potential in *Aedes aegypti* mosquitoes. *Proc Natl Acad Sci USA* 2018;115:361–366.
19. Dutra HLC, Rocha MN, Dias FBS, Mansur SB, Caragata EP, et al. *Wolbachia* blocks currently circulating Zika virus isolates in Brazilian *Aedes aegypti* mosquitoes. *Cell Host Microbe* 2016;19:771–774.
20. Rocha MN, Duarte MM, Mansur SB, Silva BDE, Pereira TN, et al. Pluriotency of *Wolbachia* against arboviruses: the case of yellow fever. *Gates Open Res* 2019;3:161.
21. Ant TH, Sinkins SP. A *Wolbachia* triple-strain infection generates self-incompatibility in *Aedes albopictus* and transmission instability in *Aedes aegypti*. *Parasit Vectors* 2018;11:295.
22. Tan CH, Wong PJ, Li MI, Yang H, Ng LC, et al. wMel limits zika and chikungunya virus infection in a Singapore *Wolbachia*-introduced *Ae. aegypti* strain, wMel-Sg. *PLoS Negl Trop Dis* 2017;11:e0005496.
23. Hoffmann AA, Iturbe-Ormaetxe I, Callahan AG, Phillips BL, Billington K, et al. Stability of the *wMel Wolbachia* infection following invasion into *Aedes aegypti* populations. *PLoS Negl Trop Dis* 2014;8:e3115.
24. Hoffmann AA, Montgomery BL, Popovic I, Iturbe-Ormaetxe I, Johnson PH, et al. Establishment of *Wolbachia* in *Aedes aegypti* populations to suppress dengue transmission. *Nature* 2011;476:454–457.
25. Freniti FD, Zakir T, Walker T, Popovic J, Pyke AT, et al. Limited dengue virus replication in field-collected *Aedes aegypti* mosquitoes infected with *Wolbachia*. *PLoS Negl Trop Dis* 2014;8:e2688.
26. Ryan PA, Turley AP, Wilson G, Hurst TP, Retzki K, et al. Establishment of wMel*Wolbachia* in *Aedes aegypti* mosquitoes and reduction of local dengue transmission in Cairns and surrounding locations in northern Queensland, Australia. *Gates Open Res* 2019;3:1547.
27. O’Neill SL, Ryan PA, Turley AP, Wilson G, Retzki K, et al. Scaled deployment of *Wolbachia* to protect the community from dengue and other *Aedes* transmitted arboviruses. *Gates Open Res* 2018;2:36.
28. Wu M, Sun LV, Vamathavan J, Riegler M, Deboy R, et al. Phylog-enomics of the reproductive parasite *Wolbachia* piipientis wMel: a
streamlined genome overrun by mobile genetic elements. PLoS Biol 2004;2:e69.

29. Huang B, Yang Q, Hoffmann AA, Ritchie SA, van den Hurk AF, et al. Wolbachia genome stability and mtDNA variants in Aedes aegypti field populations eight years after release. iScience 2020;23:101572.

30. Woollfit M, Iurbe-Ormaetxe I, Brownie JC, Walker T, Rieger M, et al. Genomic evolution of the pathogenic Wolbachia strain, wMelPop. Genome Biol Evol 2013;5:2189–2204.

31. Kent BN, Bordenstein SR. Phage WO of Wolbachia: lambda of the endosymbiont world. Trends Microbiol 2010;18:173–181.

32. Lindsey ARI, Rice DW, Bordenstein SR, Brooks AW, Bordenstein SR, et al. Evolutionary genetics of cytoplasmic incompatibility genes CIFA and CIFB in prophase wo of Wolbachia. Genome Biol Evol 2018;10:434–451.

33. LePage DP, Metcaif JA, Bordenstein SR, On J, Perlmutter JI, et al. Prophage WO genes recapitulate and enhance Wolbachia-induced cytoplasmic incompatibility. Nature 2017;543:243–247.

34. Beckmann JF, Ronau JA, Hochstrasser M. A Wolbachia debiquitinating enzyme induces cytoplasmic incompatibility. Nat Microbiol 2017;2:17007.

35. Mira A, Moran NA. Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. Microb Ecol 2002;44:137–143.

36. Wernegreen JJ. Genome evolution in bacterial endosymbionts of insects. Nat Rev Genet 2002;3:850–861.

37. Moran NA. Accelerated evolution and Muller’s ratchet in endosymbiotic bacteria. Proc Natl Acad Sci USA 1996;93:2873–2878.

38. Richardson MF, Weinert LA, Welch JJ, Linheiro RS, Magwire MM, et al. Population genomics of the Wolbachia endosymbiont in Drosophila melanogaster. PLoS Genet 2012;8:e1003129.

39. Early AM, Clark AG. Monophyly of Wolbachia pipientis genomes within Drosophila melanogaster: geographic structuring, titre variation and host effects across five populations. Mol Ecol 2013;22:5765–5778.

40. Apostol BL, Black WC, Miller BR, Reiter P, Beatty BJ. Estimation of the number of full sibling families at an oviposition site using RAPD-PCR markers: applications to the mosquito Aedes aegypti. Theor Appl Genet 1993;86:991–1000.

41. Rašić G, Filipović I, Weeks AR, Hoffmann AA. Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, Aedes aegypti. BMC Genomics 2014;15:275.

42. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 2017;3:e000132.

43. Chu J, Sadeghi S, Raymond A, Jackman SD, Nip KM, et al. BioBloom tools: fast, accurate and memory-efficient host species sequence screening using bloom filters. Bioinformatics 2014;30:3402–3404.

44. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.

45. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. Bioinformatics 2011;27:2987–2993.

46. Overmars L, van Huijsem SAFT, Siezen RJ, Francke C. CiVi: circular genome visualization with unique features to analyze sequence elements. Bioinformatics 2015;31:2867–2869.

47. Wilt A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 2012;40:11189–11201.

48. Siguiér P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 2006;34:D32–D36.

49. Varani AM, Siguiér P, Gourbeyre E, Charneau V, Chandler M. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. Genome Biol 2011;12:R30.

50. Hawkey J, Hamidian M, Wick RR, Edwards DJ, Billman-Jacobe H, et al. ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. BMC Genomics 2015;16:667.

51. Wick RR, Judd LM, Gorrie CL, Holt KE. UniCycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017;13:e1005959.

52. Darling AE, Mau B, Perna NT. Progressive Mauve: multiple alignment of genomes with gene flux and rearrangement. arXiv 2009:0910.5780.

53. Gutzwiller F, Carmo CR, Miller DE, Rice DW, Newton ILG, et al. Dynamics of Wolbachia pipientis gene expression across the Drosophila melanogaster life cycle. G3 2015;5:2843–2856.

54. Chrostek E, Marialva MSP, Esteves SS, Weinert LA, Martinez J, et al. Wolbachia variants induce differential protection to viruses in Drosophila melanogaster: a phenotypic and phylogenomic analysis. PLoS Genet 2013;9:e1003896.

55. Duarte EH, Carvalho A, López-Madrigal S, Costa J, Teixeira L. Forward genetics in Wolbachia: regulation of Wolbachia proliferation by the amplification and deletion of an additive genomic island. PLoS Genet 2021;17:e1009612.

56. Shropshire JD, Bordenstein SR. Two-by-one model of cytoplasmic incompatibility: synthetic recapitulation by transgenic expression of cifA and cifB in Drosophila. PLoS Genet 2019;15:e1008221.

57. Morrow JL, Schneider DI, Klasson L, Janitz C, Miller WJ, et al. Parallel sequencing of Wolbachia Wcr2 from donor and novel hosts reveals multiple incompatibility factors and genome stability after host transfers. Genome Biol Evol 2020;12:720–735.

58. Abbot P, Moran NA. Extremely low levels of genetic polymorphism in endosymbionts (Buchnera) of aphids (Pemphigus). Mol Ecol 2011;20:2649–2660.

59. Funk DJ, Wernegreen JJ, Moran NA. Intraspecific variation in symbiont genomes: bottlenecks and the aphid-buchnera association. Genetics 2001;157:477–489.

60. Hotopp JCD, Lin M, Madupu R, Crabtree J, Angiuoli SV, et al. Comparative genomics of emerging human ehrlichiosis agents. PLoS Genet 2006;2:e21.