Author/s:
Kriesner, P;Hoffmann, AA

Title:
Rapid spread of a Wolbachia infection that does not affect host reproduction in Drosophila simulans cage populations

Date:
2018-07-01

Citation:
Kriesner, P. & Hoffmann, A. A. (2018). Rapid spread of a Wolbachia infection that does not affect host reproduction in Drosophila simulans cage populations. Evolution, 72 (7), pp.1475-1487. https://doi.org/10.1111/evo.13506.

Persistent Link:
https://hdl.handle.net/11343/283995

Rapid spread of a *Wolbachia* infection that does not affect host reproduction in *Drosophila simulans* cage populations

Authors' names and affiliations: Peter Kriesner[1,2] and Ary A. Hoffmann[1]

[1] School of BioSciences, Bio21 Institute, The University of Melbourne, Parkville, 3010, Australia [2] E-mail: peterkriesner@yahoo.com

Corresponding author: Peter Kriesner

Short running title: Rapid spread of *Wolbachia* in *Drosophila* cages

Author Contributions: Conceived and designed the experiments: PK AAH. Performed the experiments: PK. Analyzed the data: PK. Wrote the paper: PK AAH.

Acknowledgements: Sincere thanks to Jeremy Brownlie for collecting and posting flies from Brisbane, Qld, and for helpful discussions and assistance with initial attempts to establish cage population field trials at Griffith University, Nathan campus. Also to Jason Kennington for providing flies from Perth WA, secondbite.org, and Punt Road and Oakridge Wineries, Yarra Valley, Victoria for allowing access to their properties for collection of *Drosophila*.

Funding: This research was supported by grants from the National Institutes of Health and National Health and Medical Research Council.

Data archiving: Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.n31f0t4

Abstract

Wolbachia endosymbionts that are maternally inherited can spread rapidly in host populations through inducing sterility in uninfected females, but some *Wolbachia* infections do not influence host reproduction yet still persist. These infections are particularly interesting because they likely represent mutualistic endosymbionts, spreading by increasing host fitness. Here we document such a spread in the *w*Au infection of *Drosophila simulans*. By establishing multiple replicate cage populations, we show that *w*Au consistently increased from an intermediate frequency to near fixation, representing an estimated fitness advantage of around 20% for infected females. The effective population size in the cages was estimated from SNP markers to be around a few thousand individuals, precluding large effects of genetic drift in the populations. The exact reasons for the fitness advantage are unclear but viral protection and nutritional benefits are two possibilities.

Key words: Wolbachia, fitness advantage, rapid spread, host fitness, cage populations, viral protection

Introduction

Wolbachia bacterial endosymbionts have considerably streamlined genomes and reduced metabolic capabilities in comparison with typical free-living bacteria (Wu et al. 2004; Klasson et al. 2009; Sutton et al. 2014), and infection is maintained across host generations by strict maternal inheritance. It is therefore necessary that within given host populations, *Wolbachia*-infected females gain some form of advantage relative to their uninfected counterparts as well as reliably transmitting the infection to their offspring. *Wolbachia* can enhance their own spread by inducing a range of host reproductive pathologies. The most prevalent of these known as cytoplasmic incompatibility (CI), is now thought to entail a type of toxin-antitoxin mechanism whereby developing sperm are modified by a factor secreted by CI-inducing *Wolbachia* present in the germline cells of infected males (Beckmann et al.

2017; LePage et al. 2017). When ova bearing the same strain of *Wolbachia* are fertilized by such modified sperm, a *Wolbachia*-derived complementary binding factor effects rescue of the modification which otherwise leads to increased embryonic mortality. This results in a selective advantage for infected females as the infected ova they produce can be fertilized successfully by sperm from any male.

However, this type of advantage is frequency-dependent. For host populations containing only a few infected males, the overall effect of *Wolbachia*-induced CI may be small. Because vertically-transmitted intracellular bacteria such as *Wolbachia* depend entirely on host-provided resources, it is expected infected hosts will incur at least some underlying fitness cost. In the absence of some other form of compensating benefit, mathematical models (Caspari and Watson 1959; Fine 1978; Hoffmann et al. 1990; Turelli 2010) predict CI-inducing *Wolbachia* infection will exhibit bistable dynamics with a tendency to be eliminated from host populations where frequency is below an unstable equilibrium. The utility of these models is supported by observed dynamics for *Aedes aegypti* populations at field sites in North Queensland, Australia where, more than two years after large scale releases of laboratory-reared mosquitoes ended, *w*Mel strain infection remains at high frequency in the initial release areas (Hoffmann et al. 2014) despite infected individuals exhibiting a fitness disadvantage. However the infection has not spread effectively across surrounding geographic barriers—apparently because the few infected individuals that manage to migrate across such barriers are too few in number to generate population infection frequencies above the inferred unstable point (Barton and Turelli 2011; Hoffmann et al. 2011; Hoffmann et al. 2014).

The persistence and spread of infection is particularly challenging for *Wolbachia* strains which induce very weak or no apparent reproductive manipulation (e.g. Hoffmann et al.

1996; Zhang et al. 2010; Zhao et al. 2013; Hamm et al. 2014); these infections persist in field populations even with imperfect maternal transmission, and they have also likely spread from initially very low frequency (Kriesner et al. 2013) following presumably very rare horizontal transfer. Such strains may arise due to an evolutionary conflict between *Wolbachia* and their hosts with the expectation that either host resistance, or a more mutualist relationship will emerge, at least over evolutionary timescales (Turelli 1994; Herre et al. 1999). There is empirical evidence from *Wolbachia* for both these phenomena: when transinfected into novel hosts *Drosophila simulans* (Poinsot et al. 1998) and *Ae. aegypti* (Walker et al. 2011) the *w*Mel strain induces almost complete CI but the effect is weak to moderate and generally detected only for young males in its native host, *D. melanogaster* due to host background effects (Reynolds and Hoffmann 2002). Very recent evolution of host resistance to *Wolbachia*-induced male-killing has been demonstrated for populations of the tropical butterfly *Hypolimnas bolina* in Southeast Asia (Hornett et al. 2006) and even documented in action in Samoa (Charlat et al. 2007). Furthermore, the *w*Ri strain which had previously invaded *D. simulans* populations in California (Turelli and Hoffmann 1991) changed over the following 20 years from inducing a significant fecundity deficit to providing a fecundity boost (Weeks et al. 2007).

Provision of host fitness benefits that are frequency-independent may therefore be necessary to explain *Wolbachia's* evolutionary success. These have been hypothesized for some time in *Drosophila* (Hoffmann et al. 1998; Harcombe and Hoffmann 2004). Candidates include protection against natural enemies (cf. Fenton et al. 2011), in particular anti-viral protection (Teixeira et al. 2008; Osborne et al. 2009), or metabolic enhancement (Brownlie et al. 2009; Hosokawa et al. 2010; Moriyama et al. 2015), either of which may be context dependent (Jaenike 2012).

Despite this, few studies have directly investigated fitness enhancement for insect hosts infected with *Wolbachia* or other endosymbionts, and even fewer have demonstrated them under field conditions. Fast et al. (2011) identified a very substantial *Wolbachia* effect on fecundity in laboratory populations of *D. mauritiana*. However increased egg production alone may not translate to greater numbers of offspring surviving to adulthood. Himler et al. (2011) and Hamm et al. (2014) inferred sizeable positive host fitness effects—for a strain of Rickettsia in populations of *Bemisia tabaci* and the *w*Suz strain of *Wolbachia* in populations of *D. suzukii* respectively—to explain rates of infection spread determined from observed field data, and in the latter case, infection persistence over time despite substantial maternal transmission leakage.

Most experimental manipulations to measure *Wolbachia* effects involve situations where infection dynamics are also driven by reproductive manipulations (e.g. Dobson et al. 2002; Axford et al. 2016). While these experiments provide convincing tests of the impact of reproductive manipulations associated with *Wolbachia*, they cannot detect subtle fitness effects. Stochastic drift could lead to a CI-inducing *Wolbachia* infection—even one which imposes a net fitness cost—increasing in frequency in a host population to beyond the relevant unstable point and then becoming established at a higher stable equilibrium. Although stochastic drift alone may be an unlikely mechanism for infection to become established in field populations of any significant size from initial invasion (Egas et al. 2002; Jansen et al. 2008), it may well be relevant in smaller populations such as those typical of laboratory setups and for host effects that are positive but moderate in size.

In *Drosophila simulans*, the most well-known and investigated *Wolbachia* endosymbiont is the *w*Ri infection (Turelli and Hoffmann 1991, 1995), but populations of this species can also be infected by other *Wolbachia* variants (Ballard 2004) including the *w*Au strain which does

not induce CI (Hoffmann et al. 1996). Kriesner et al. (2013) showed that *w*Au had spread in eastern Australian populations of *D. simulans* before being displaced by *w*Ri, and they were able to infer a fitness advantage for *w*Au-infected flies relative to uninfecteds (F) on the order of $F = 1.06$ or more, and a greater fitness advantage for *w*Ri-infected flies of $F \geq 1.08$. This suggests that positive fitness effects of *w*Au should be detectable in populations of a sufficient size. Here we therefore directly assess the comparative fitness of infected and uninfected *D. simulans* hosts for both the *w*Au and *w*Ri infections in large cage populations using population modelling to inform the experimental design.

Wolbachia-mediated protection of hosts has been most extensively demonstrated in laboratory settings against viral infection, with *w*Au in particular having been found to provide highly effective virus blocking (Martinez et al. 2014; Ant et al. 2018). We thus also explore as a secondary aim the hypothesis that these *Wolbachia* strains influence host fitness by inhibiting pathogens—primarily viruses—that are potential afflictions of *D. simulans* in the field.

Materials and Methods

Simulations

To investigate the potential impact of stochastic events on local *Wolbachia* infection dynamics in constrained small- to moderate-sized populations, an established model (Hoffmann and Turelli 1997) which allows for set levels of maternal transmission loss (μ) and hatch rate from notionally incompatible crosses (H) was modified to incorporate randomly variable outcomes for these parameters, and for male mating success, crossing types arising from matings, female reproductive success, and chance of resulting viable embryos surviving to adulthood. Initial generations for population trials comprised set numbers of adult females with up to 3 infection status types (representing uninfected, *w*Au-

infected, or *w*Ri-infected). It was assumed these females had pre-mated with a male of the same infection type. Although the model could accommodate 3 way comparisons, most trial runs only compared uninfecteds with one of the infected types.

Monte Carlo simulations and randomization of population events were enacted using functions of the PopTools package (Hood 2011) in a spreadsheet model for discrete generations with outcomes rounded to the nearest integer where appropriate and the following parameters:

- Number of reproductively successful females in each cohort ($N_{CS♀}$) where cohort is based on infection status: random binomial with n = total number of females comprising cohort ($N_{C♀}$) and $p = 0.92$.
- Number of ova produced for each cohort: $N_{CS♀} \times F \times \omega$ (rounded to nearest integer) where F = relative fitness of cohort and ω = normally distributed random variable with mean = 64.9 and $\sigma = 23.5/\sqrt{N_{CS♀}}$. Values for p, mean of normal distribution and σ used here were based on data obtained from fecundity assays performed previously using outbred *D. simulans* laboratory populations of varying infection status (data not shown).
- Number of uninfected ova produced by infected mothers per cohort: random binomial with n = total number of ova produced per cohort and p = rate of maternal transmission leakage (μ).
- Number of uninfected ova fertilized by sperm from *w*Ri-infected males and rendered inviable due to CI: random binomial with n = number of uninfected ova arising per cohort and $p = 1 - H$ (H being the hatch rate from notionally incompatible crosses).

- Number of viable ova surviving to adulthood per cohort ($N_{C,t+1}$): random binomial with n = total viable ova produced per cohort and $p = K / N_{VO}$ (proportionally adjusted so that total adults in new generation ≤ K), where K = carrying capacity of the environment (maximum allowable number of adults) and N_{VO} = total viable ova produced for all cohorts.
- For each new cohort arising from each new generation of adults, the number that are female ($N_{C♀,t+1}$): random binomial where $n = N_{C,t+1}$ and $p = 0.5$.
- For each new cohort; the number of males ($N_{C♂,t+1}$): $N_{C,t+1}$ - $N_{C♀,t+1}$.
- Number of successful male matings for each new cohort: random drawn from Poisson distribution with mean = $N_{C♂,t+1}$, proportionally adjusted so that total successful male matings for all cohorts = total successfully reproducing females for all cohorts.
- Number of reproductively successful females in each new cohort ($N_{CS♀,t+1}$) that mated with males of each respective infection type: random binomial with $n = N_{CS♀,t+1}$ and p = proportion of total successful male matings involving a male of that infection type.

Simulations for populations initiated with pre-mated females only were performed for discrete host generations with a mean value of $H = 0.55$ for *w*Ri based on previously-obtained estimates (Carrington et al. 2011) and a nominal mean value of $\mu = 0.003$ for both strains. Previous estimates for μ of 0.048 for *w*Ri infection (Carrington et al. 2011) and 0.023 for *w*Au infection (Kriesner et al. 2013) have been derived for field populations but these data were highly variable with only a few females producing most of the leakage in both cases. Factors giving rise to maternal leakage in field populations are unclear but it does not occur in laboratory populations (Hoffmann et al. 1990). Mating events for F_1 and subsequent generations were assumed to be random as there is little evidence for assortative mating

based on *Wolbachia* infection status among wild type *D. simulans* or *D. melanogaster* (Champion de Crespigny and Wedell 2007) or indeed other insects (although see Vala et al. 2004). Male mating success was assumed to be Poisson distributed as suggested by findings of Joshi et al. (1999).

Trials were run varying initial population size (N_i), initial infection rate (p_i), carrying capacity (K) of the environment (sensu Hui 2006)—modelled as the maximum number of potentially mating adults, and relative fitness (F). Given the high reproductive rate of *Drosophila* females (built into the model), adult census population size (N) = K was expected for F_1 and subsequent generations and model outcomes were relatively insensitive to small variations in N_i. Monte Carlo analysis was employed to identify parameters that would enable discrimination with reasonable confidence between infection dynamics driven by positive fitness effects—using F equivalent to minimum values of 1.06 for *w*Au and 1.08 for *w*Ri as previously inferred for field populations in eastern Australian (Kriesner et al. 2013)—and dynamics driven only by CI (in the case of *w*Ri) and stochastic drift (i.e. F = 1). For *w*Ri trials, this required quite low values for p_i.

Experimental populations and cages

Preliminary experiments were undertaken to estimate larval populations that could be supported by laboratory medium, based on larval stages of *D. simulans* only utilizing the upper few millimetres of food medium provided. Conical laboratory stock bottles containing ~40 ml of food medium and with ~205 cm^2 internal surface area including food medium surface of ~ 21.2 cm^2 typically support *Drosophila* populations on the order of 250 to 300 adult individuals. On this scale containers with 510 cm^2 food medium surface and 0.5 m^2 total internal surface area should support populations of >6,000 adults.

Cages were constructed using archive storage box size polypropylene containers with dimensions of 40 × 32 × 26 cm, internal volume of ~33 litres, internal surface area of ~0.63 m^2 and sealed lids. Holes were cut in the top and long sides of the boxes, and marine grade stainless steel fine woven wire mesh with nominal aperture of 0.132 mm, wire diameter 0.08 mm and 39% open area was taped and plastic welded in place covering the holes to provide ventilation. The nominal aperture opening of the wire mesh was chosen to be smaller than the average *D. simulans* egg width of 0.18 mm (Markow et al. 2009).

Internal access was provided via a double layer of tubular elastic support bandage material (Tubigrip, Mölnlycke Health Care, Sweden) attached to an opening at one end of each cage. The elastic sleeve was knotted and additionally sealed with a bolted-on plastic cover plate when not in use. Laboratory food medium consisting of dextrose (7.5% w/v), cornmeal (7.3% w/v), inactive yeast (3.5% w/v), soy flour (2% w/v), agar (0.6% w/v), Nipagin (Methyl 4-hydroxybenzoate, 1.6%) and acid mix (1.0% 10:1 propionic acid: orthophosphoric acid) was provided in polypropylene trays with resealable clip-on lids and external dimensions of 17 × 11 × 7 cm. Four trays each containing ~300 ml of food medium with 150 cm^2 surface area were provided for each experimental cage initially. Subsequently, for each cage every 10 or 11 days, the oldest two food medium trays were removed and replaced with new trays, the medium in the remaining two trays was mixed to expose unused surface area, and a freshly peeled banana (skin previously unbroken) was added to one of the remaining trays. Additionally, paper index cards were placed in new food medium trays and the sides of all trays were roughened with sandpaper to facilitate pupation. Trays containing prepared food medium were stored at 14° – 16°C for up to one month prior to use.

Experimental cages were housed together in a heated and well-insulated storage room with natural daylight and loosely enclosed in plastic to prevent desiccation stress while still

allowing ventilation. Air temperature and humidity inside one representative cage was monitored with a data logger (Hygrochron DS1923, iButtonLink Technology, WI). Temperature was consistently maintained within the range of 24°C – 26°C and humidity within ~60% – 70% RH for the duration of the experiments. All cages were stored on separate plastic trays.

Stocks

Three separate laboratory lines of *D. simulans* were established, each from combined offspring of several individual isofemale lines of the same *Wolbachia* infection status, with each isofemale line having previously been confirmed to be either *w*Au-infected, *w*Ri-infected or uninfected respectively. The *w*Au-infected and uninfected lines were initiated with field-caught females from Perth, Western Australia, while the *w*Ri-infected lines were initiated using field-caught females from Brisbane, Queensland.

For the *w*Au-infected and uninfected laboratory lines, around 100 virgin females from each line were outcrossed with ~100 field-caught males collected from Brisbane for 5 generations and then with ~100 field-caught males collected from Burnley, Victoria for a further 2 generations. For the *w*Ri-infected laboratory line, ~100 virgin females were initially outcrossed once with ~100 males from the uninfected laboratory line and then outcrossed with field-caught males from Brisbane and Burnley as described above. This resulted in outbred *w*Au-infected, *w*Ri-infected, and uninfected mass bred lines. Each mass bred line was maintained at low density in 250 ml capacity glass bottles with standard fly food medium for at least 2 further generations. Overall numbers of flies were increased over this period to provide sufficient numbers of females for the trial cage experiments. Each mass bred line was then retested to confirm infection status (n = 15 per line) prior to use in the trials. Females used to establish the trial populations were lightly anaesthetized, counted and allowed to

recover before being introduced into the cages with a temporary food source. Females lacking hardened cuticles and mature pigmentation (i.e. potentially less likely to have already mated) were excluded. The *w*Au type trials were initiated with 350 females from the *w*Au-infected mass bred line and 650 females from the uninfected mass bred line, whereas *w*Ri type trials were initiated with 60 females from the *w*Ri-infected mass bred line and 1440 females from the uninfected line. After a further 12 hrs recovery, the temporary food source was removed and replaced with four fresh trays of food medium as described above.

Cage conditioning and initiation

Because *Wolbachia* effects on host fitness may be mediated by exposure to other microbes, we were interested in comparing frequency changes in cages which might have different levels of microbe activity—in particular where this entails oral-faecal transmission. In an attempt to generate different levels of microbe activity, one set of trial cages (herein termed 'high exposure' cages) was pre-conditioned by allowing prior surface contact with flies collected from a large and long-established field population from Burnley, Victoria, alongside a set of ('low exposure') control cages that were initially clean and not preconditioned. It was hypothesized that at least some pathogens that potentially commonly infect *D. simulans* in the field are typically enterically-acquired and rely on horizontal transmission via faeces or oral excretion. This has been demonstrated for Nora virus (Habayeb et al. 2009) although knowledge regarding the mode of transmission for other such pathogens is lacking.

Just prior to commencement of the experiment trials, several thousand adult flies were collected from the field sites and introduced in approximately equal numbers into the high exposure cages (Table 1), along with a temporary source of food medium. These cages were maintained under moderate humidity to prevent desiccation for 24 hrs. The field flies (and

food) were then removed, cage lids resealed, and the experimental populations subsequently introduced within ~2 hrs.

Populations polymorphic for either *w*Au or *w*Ri infection were initiated in the high or low exposure cages (Table 1) by introducing a set number and ratio of *D. simulans* pre-mated adult females sourced from three separate mass bred laboratory lines of previously confirmed infection status (*w*Au-infected, *w*Ri-infected and uninfected respectively).

Four separate trays containing food medium were introduced into each cage at the commencement of the trials. The trays, which together occupied most of the floor area of the cages, were then rotated with half of them being replaced every 10 or 11 days. The cages were maintained at 24° – 26°C. As the development time for *D. simulans* laboratory populations at 25°C can be as short as 8.3 days with low population density (Woods et al. 2002) but likely to be somewhat longer for *Drosophila* generally at higher densities (Ashburner et al. 2005), it was reasoned that a 10 to 11 day food medium replenishment cycle would facilitate population turnover. It was expected that fresh medium would be quickly utilized by adult females for oviposition, and that larvae hatching early from newly deposited eggs would generally outcompete later developing smaller instars and reach adult stage within one replenishment cycle. Observations of large numbers of pupae developing in food medium trays after a few days, and of large adult population sizes in cages after the first 10 days, support these conjectures.

Sample collections and *Wolbachia* infection status

Adult flies sampled from within the experimental cages were collected using a vacuum filter method with the aim of reducing bias in favour of sampling the healthiest, most vigorously flying individuals. Samples for subsequent PCR assays were preserved in 100% EtOH and

stored at -20°C until use. Samples for RT-PCR assays were returned to the laboratory in vials at low density, snap frozen in LN_2, and stored at -80°C until use in PCR.

Sampled flies were assayed for *Wolbachia* infection status and strain type using a real-time PCR / high resolution melt (RT-PCR/HRM) method designed to amplify a fragment of the wsp gene (Lee et al. 2012) as previously described in Kriesner et al. (2013). DNA extraction for individual flies was carried out using a standard Chelex based method. Separate *D. simulans* specific primers (Dsim_RpS6_Fwd: 5'-CCAGATCGCTTCCAAGGAGGCTGCT-3'; Dsim_RpS6_Rev: 5'-GCCTCCTCGCGCTTGGCCTTAGAT-3') were used as an internal control to test for the presence of DNA and correct PCR conditions. Individual flies from both the *w*Au- and *w*Ri-infected *D. simulans* mass bred laboratory lines were used as positive and strain type controls for each PCR plate while individual flies from the uninfected mass bred line and no template PCR reactions were used as negative controls. We also screened flies sampled from half of the experimental cages for the presence of selected microbes as described in Methods S1.

Statistical tests were performed using R version 3.3 (2017).

DNA library preparation for ddRADseq

Genomic DNA extractions were performed for batches of 19 all male or all female *D. simulans* adults collected at 3 separate time points (April, July and November) from each of the 8 trial cages using High Pure PCR Template Preparation Kits (Roche Applied Science, Mannheim, Germany) in accordance with the manufacturer's instructions with the additional step of homogenizing the sample material in a mini-beadbeater with 2 × 3 mm glass beads per batch for 2 × 2 mins at 25 Hz. The amount of genomic DNA was quantified for each DNA extraction using a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad CA, USA). A single reduced representation ddRADseq library of pooled DNA samples was produced by

simultaneous digestion with two restriction enzymes, ligation of combinatorial in-line barcoded adapters, size selection of fragmented DNA, and PCR amplification such that only the subset of restriction digest fragments generated by cuts with both restriction enzymes (i.e. fragments with different ends) would be sequenced. The protocol was a modified version of the protocol described in Peterson et al. (2012) with in-line DNA barcode sequences included in P2 as well as P1 type adapters but with a common set of PCR primers. The complete library comprised 76 individuals (2 batches of 19 males + 2 batches of 19 females) for each time point from each of the cages, uniquely barcoded, in replicate (i.e. 8 × 3 × 2 = 48 barcode combinations, total size of pools across both technical replicates = 152 individuals).

To inform the choice of enzymes and appropriate size selection range, in silico digestions of the *Drosophila simulans* reference genome r2.02 (flybase.org) were performed with the ddSilico utility program (Rašić et al. 2014) to calculate potential ddRAD loci numbers. From each batch of 19 individuals, 500 ng of extracted genomic DNA was digested with 2 units of SphI-HF and 10 units of MluCI restriction enzymes (New England Biolabs, Ipswich MA, USA), 10 µg BSA, 1× NEB CutSmart Buffer and nuclease-free water in a 50 µl reaction for 3 hrs at 37°C without a heat kill step. The digestion products were cleaned with 1.5× volume of carboxylate-modified paramagnetic beads solution (Sera-Mag SpeedBeads, Thermo Fisher Scientific, Waltham MA, USA).

Stock solutions for 12 × P1 type and 4 × P2 type in-line barcoded adapters were created as required by annealing 10 µl each of complimentary 1.1 and 1.2, or 2.1 and 2.2 type oligonucleotides (see Table S1 for oligonucleotide sequences) at 40 µM final concentration in 1× annealing buffer (10 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA) and 50 µl reaction volume in a thermocycler with initial temperature of 97.5°C for 150 secs, decreased by 2.5°C every 1 min for 31 cycles, then stored at 4°C until use. Stock solutions of adapters were

further diluted with 1× annealing buffer based on output of the ddRAD ligation molarity calculator developed by (Peterson et al. 2012) to form adapter working stocks. For each technical replicate, equimolar amounts of digested and cleaned genomic DNA extracted from batches of flies comprising the same sampling pool were then combined and distinctively barcoded by ligating a unique combination of P1 and P2 type adapters (i.e. 2 × 24 barcode combinations). Ligation reactions with 2 µl of 1.77 µM P1 and 2 µl of 5.9 µM P2 adapter working stocks, 1000 units of T4 ligase and 1× T4 buffer (New England Biolabs) in a 45 µl reaction volume were incubated at 16°C overnight. Equimolar amounts of adapter ligated DNA from all pools were then combined and cleaned with 1.5× volume paramagnetic beads solution and eluted in 60 µl nuclease-free water. Size selection of fragments between 340–600 bp was performed using a Pippin Prep 2% gel cassette (Sage Sciences, Beverly MA, USA). Size selected DNA recovered from the Pippin Prep cassette was then used as a template in 15 µl PCR reactions, each with 7.5 µl of Phusion High Fidelity 2× Master Mix (New England Biolabs), 3 µl each of 10 µM PCR 1 and PCR 2 primers (see Table S1 for PCR primer sequences), and 1.5 µl DNA template with PCR conditions: 98°C for 30 s; 12 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 70 s, and final elongation at 72°C for 5 min. PCR products from eight such reactions were then combined and cleaned with 0.8× volume paramagnetic beads solution to make the final library. Sequencing was performed at the Australian Genome Research Facility, Parkville, Australia on a single lane of an Illumina HiSeq2000 platform run.

Bioinformatic data processing

Raw sequencing reads were checked for quality using FastQC (Andrews 2010). Initial R1 and R2.fastq Illumina files were then demultiplexed using the process_radtags program from the Stacks pipeline (Catchen et al. 2013) with default parameters and allowing for 1 mismatch in

the barcode-restriction enzyme cut site sequence, resulting in pairs of R1 and R2.fastq files for each of the 24 pools sampled from the trial cages in two technical replicates. After remainder reads were discarded, the first technical replicate included 54.1% of the sequence data with the second replicate comprising the balance. At this point both R1 and R2.fastq files belonging to the same pools of fly samples across the two replicates were merged with read coverage across pools varying by no more than 21% from the average coverage for the 24 pools.

Read pairs were then aligned to the reference genome, *Drosophila simulans* r2.02 with the Bowtie 2 aligner (Langmead and Salzberg 2012) and settings of '-X 750 -D 20 --no-discordant --no-overlap --no-mixed'. The resulting files were sorted with the Picard v. 2.7.1 (Broad Institute 2014) SortSam tool, and converted to .bam format and filtered for mapping quality using the Samtools v. 1.3.1 (Li et al. 2009) view command with setting '-q 20', and indexed using the Samtools index command. Multi-pileup files were created using the Samtools mpileup command to combine the .bam file data for pooled samples collected at the April, July and November time points for each of the eight experimental trial cages. Each mpileup file was filtered to exclude genomic regions with coverage of ≤ 10× and > 600× using the Unix awk command, and converted to sync format with the mpileup2sync.jar tool in Popoolation2 (Kofler et al. 2011).

Sync files were read using the read.sync function from the Nest package (Jónás et al. 2016) in R. Estimates of N_e were then produced using the estimateNe function from the Nest package with arguments of 't=22, ploidy=2, truncAF=0.04, method=c("P.planI", "JR.planI", "W.planI"), Ncensus=5000, poolSize=c(152,152)' for allele frequency estimates over the April to November time interval, or 't=11' for either the April to July or July to November intervals.

Results

Stochastic modelling of infection dynamics

Across 40 host generations, trial outcomes were highly variable for K on the order of 500 to 1500 (Figure 1A-B), but more consistent for $K \approx 5000$ or greater. For $K = 6000$, there was no intersection of 95% confidence limits derived from Monte Carlo simulations for the F_{10} and subsequent generations from *w*Au trials (Figure 1C), or for F_{21} and subsequent generations for *w*Ri trials (Figure 1D). If higher values for μ were used, trials using positive values for F diverged further on average from trials with $F = 1$. Based on these simulations, we surmised that experimental populations with an adult census size (N) consistently maintained at > 6000 for ~ 20 to 30 host generations or longer could reasonably be expected to yield infection dynamics predominantly driven deterministically.

Population cage infection frequencies

Adult samples were collected from each of eight experimental cages designed to house > 6000 adult flies at intervals of approximately 3½ months; following the 10th (April samples) and 21st (July samples) food tray rotations, and at the conclusion of the trials (10½ months after commencement)—around 10 days after the final (31st) rotation (November samples). These intervals were estimated to approximately correspond with the passage of 10, 21 and 32 host generations respectively. Whilst completion of up to 38 to 39 host generations may have been theoretically possible for some proportion of the cage populations by the end of the trials, this would only be expected if the maximum estimated low density development rate had been maintained throughout. Conversely, if generation times were more closely aligned with the food medium replenishment cycle, at least some adult flies would be expected to persist beyond each cycle resulting in somewhat longer average generation times.

Individual flies from the sample collections were subsequently assayed for species, infection status, and strain type by PCR. Despite the long duration of the trials, there was no evidence for any of the cage populations having been compromised. All samples assayed as infected were of the expected strain type, and no *D. melanogaster* males or individuals of other fly species were found in any cage at any point despite being intermittently present in the surrounding area.

Results from these assays, summarised in Figure 2, generally indicate a large and faster-than-expected increase in infection prevalence: for all cages initially polymorphic for *w*Au (Figure 2A), infection frequency had increased significantly by the first sampling period (exact binomial test, $P < 2.2 \times 10^{-16}$) and apparently reached fixation by the end of the trials (all samples from the final collection assayed were found to be infected). The pre-conditioning treatment applied to 'high exposure' cages did not have a consistent effect on subsequently assayed infection frequencies. The observed infection frequency for one of the control replicates (cage G) for the April samples tested was similar to that for the two 'high exposure' replicates (cages A and C), but it was significantly lower (Chi-squared test, $P = 1.2 \times 10^{-5}$) for the second control replicate (cage E). This difference persisted with infection frequency observed for cage E also significantly lower than for the other three *w*Au cages for the July samples tested ($P = 0.0053$).

An increase in infection frequency was also detected for cages initially containing some *w*Ri-infected females (Figure 2B) with the exception of cage F (all samples assayed from this cage were uninfected). The apparent loss of *w*Ri infection in cage F is unexplained. For the other trial cages where *w*Ri persisted, infection frequency for the April samples was significantly lower for cage H than for the two 'high exposure' cages B and D (Chi-squared test, $P < 2.2 \times 10^{-16}$). However, dynamics of infections causing CI are expected to be sigmoidal in shape and

inherently more diverged around the maximum rate of increase than expected for simple linear dynamics. For the July samples tested, infection frequency was at consistently very high levels for all three cages.

Further Monte Carlo simulations were run to examine parameters that would best fit the observed data, again assuming $N = 6000$, and that the sampling time points equated to 10, 21 and 32 elapsed host generations. The absence of uninfected individuals among the final sampling collection implies no or negligible maternal transmission loss (μ) within the cage environments. For *w*Au with $\mu = 0$, the rates of infection frequency increase observed are best explained by relative fitness of infected flies (F) being on the order of 1.19 – 1.22× that of uninfecteds. However no single constant value for F entirely matches the observed data for all three sampling time points. It is unlikely imperfect transmission would have gone undetected if actual μ was $> \approx 0.002$ although some minimal rate of leakage would suggest a slightly higher range of plausible values for F.

Dynamics for *w*Ri are more complex and more variable. With $\mu = 0$ and complete CI ($H = 0$), simulations with F on the order of 1.14 or greater fit observations from cages B, D and H for the first sampling time point considered in isolation, but F on the order of 1.20 – 1.27 with modest leakage ($\mu = 0.01$) and partial CI ($H = 40\%$) better explains the overall observations. Note that an underlying modest rate of transmission leakage may be less apparent for *w*Ri because when infection frequency is relatively high newly arising uninfected ova will be continually suppressed by CI.

As we were not able to accurately determine the proportion of older flies which persisted in trial populations beyond each food medium replenishment cycle, the actual number of host generations elapsed at each of the sampling time points may have been somewhat fewer on average than the assumptions used in these simulations. This would imply F of greater

magnitude. Alternatively, *F* may not have remained constant across the duration of the experiments.

We also detected the presence, at low frequency, of some viruses (Kallithea, Nora and La Jolla viruses, but not Thika virus) that had been previously identified as reasonably common infections among *D. simulans* field populations (Webster et al. 2015) from individual flies sampled from our population cages. These results are summarized in Table S2. We did not find an association between *Wolbachia* infection status and presence of any particular virus. However low prevalence of these specific viruses, combined with relatively few *Wolbachia*-uninfected flies among the samples collected, resulted in limited statistical power to detect any such association.

Effective population size

Cages were designed with sufficient internal volume, surface area and available food medium to accommodate large populations based on scaled-up requirements for typical *Drosophila* laboratory stocks. Several lines of evidence support the premise that a large value for *N* on the order of 6000 or more adult flies was maintained in the experimental cages over the course of the trials: founder populations comprised a set number of individuals and visual estimates of *N* from regular observations at intervals of 10 to 11 days or less indicated a population size several times as large as the apparent size of founder populations in all cages. There was however noticeable fluctuation in apparent *N* over time for most of the cages. An estimate for *N* was obtained at the conclusion of the trials for three of the cages by collecting all remaining live adults, counting and weighing a proportion of them, and extrapolating an estimate for *N* by weighing the rest that were not counted directly. This yielded values for *N* of ~8820 for cage C, ~9310 for cage E and ~6960 for cage D. Note that such an estimate did not include the substantial numbers of other life stage individuals present in the cages.

To test for bottlenecks influencing effective population sizes, a double digest Restriction-site Associated DNA sequencing (ddRADseq) approach (Peterson et al. 2012) was used to obtain SNP allele frequency data with pooled samples of additional adults which were collected from each of the eight trial cages at the same three time points as samples used for individual PCR assays. Each of these pools comprised 152 adults. Sequencing of the resulting DNA library yielded 200 million Illumina 100 bp paired-end reads with acceptably even coverage across the 24 separately barcoded pools (std. dev. = 10.7%). Pooled sequencing was used because, with the same level of finite resources, it is expected to outperform individual-based sequencing for estimating allele frequencies (Schlötterer et al. 2014).

Estimates for short-term effective population size (N_e) were then derived using temporal changes in allele frequencies across the three sampling time points. As estimating N_e from pooled sequence data requires two phases of sampling - individuals from whole populations, then sequencing reads from a pooled DNA library - an estimator recently developed by Jónás et al. (2016) and implemented as an R package ('Nest' version 1.1.8) was used to account for the additional variance component introduced by the use of pooling. As the individuals comprising the sample pools were collected a few days after a rotation of food medium trays, it was assumed those pools had been sampled after reproduction in accordance with the first of two sampling schemes originally proposed by Nei and Tajima (1981). Alternative estimates for N_e were also derived using two other established and commonly used methods originally developed by Waples (1989), and Jorde and Ryman (2007) that are also implemented with the 'Nest' ver. 1.1.8 package. However neither of these estimators explicitly addresses the variance component arising due to pooled sequencing (Jónás et al. 2016). Each of the estimates derived assumed $N = 5000$ and excluded SNPs with minor allele frequency < 0.04. Although the Jónás et al. (2016) method notionally requires prior knowledge of N, the estimates for N_e derived with this method were relatively insensitive to

varying N based on testing the effect of a range of N values both higher and lower than 5000 (data not shown).

Estimates of N_e calculated with each of the three methods, excluding genomic regions with coverage ≤ 10× and > 600×, and based on SNP allele frequency changes between the first and last sampling periods (assuming 22 generations between measurements), are shown in Table 2. The very large negative result for cage 'G' here, and negative estimates in this context in general, can be interpreted as no evidence of genetic drift, with N_e being essentially infinite (Peel et al. 2013). This is plausible where the number of generations between estimates is small—as was the case for these trials—and actual N_e is relatively large, such that drift has not yet had a significant effect on allele frequencies (Jónás et al. 2016).

These results, together with other estimates of N_e (Table S3) obtained for the shorter time intervals between the April and July sampling time points, and the July and November time points (assuming 11 generations each), and when considering each of the major chromosome arms of the *D. simulans* genome in isolation (assuming 22 generations), show a consistent pattern where most estimates derived with the Jónás et al. (2016) model are on the order of one to a few thousand with some, seemingly asymptotically trending, very large negative or positive results. The lowest estimates were derived for the cage 'H' population. Although the Jorde and Ryman (2007) method gave smaller estimates, simulations performed with pooled sequencing data by Jónás et al. (2016) found this method had a downward bias, particularly for allele frequency estimates made only a few generations apart, due to additional sampling variance.

Taken together, the estimates derived for N_e suggest that the trial populations were not subject to substantial population bottlenecks. In line with assumptions for the stochastic modelling, if most females are fertile and male mating success is Poisson distributed, the

effective size of a given population should be no more than ~ 0.75× that of the census size. Accordingly, the estimates obtained for N_e here are broadly consistent with expectations for N. However, for populations comprising thousands of individuals, existing methods with data from pooled sequencing have inherently limited statistical power to provide more precise estimates of N_e over short time intervals.

Discussion

Estimates of N_e for the *D. simulans* experimental populations derived from temporal changes in SNP allele frequencies support the premise that stochastic fluctuations did not have much impact on changes in *Wolbachia* infection frequencies. Large and faster-than-expected increases in infection frequency for all four of the *w*Au cages argues for a substantial fitness advantage for *Wolbachia*-infected flies in the physically-confined environment of these trials. The results from three of the four *w*Ri trial cages are consistent with a similar or slightly greater fitness advantage for that strain. This is evidence for positive fitness effects separate to any host reproductive manipulation for *Wolbachia* in *Drosophila*.

The magnitude of the relative fitness parameter (*F)* derived from model simulations required to explain the observed *Wolbachia* frequency increases is at least double, and up to 4× as large as inferred previously in field populations of *D. simulans* from eastern Australia for both *w*Au and *w*Ri (Kriesner et al. 2013). This suggests that one or more elements of the environment encountered by flies in the trial cages differed from the average environment experienced by flies generally in field conditions in a way which increased the fitness difference between *Wolbachia*-infected and uninfected flies. Field populations experience

fluctuating humidity and temperature. Variations in temperature can affect the relative costs or benefits associated with endosymbiont infection generally (Corbin et al. 2017). Fluctuating high temperatures can reduce the density of *Wolbachia* in infected hosts (Ross et al. 2017) and temperature low enough to induce reproductive dormancy can be associated with increased fitness costs for *Wolbachia* infection in *D. melanogaster* (Kriesner et al. 2013). The experimental populations in this study were provided with a standard laboratory fly food medium supplemented with banana unlikely to have been deficient in any key nutrient. The trial cage environment fostered intense competition among developing larvae for available food resources that likely placed a premium on developmental speed. If *Wolbachia* can help promote efficient host metabolism as has been suggested (Brownlie et al. 2009; Moriyama et al. 2015) then this effect may have been more pronounced in the trial cage environment system.

Additionally, flies infected with *Wolbachia* may have been protected to some degree against deleterious effects of one or more pathogens present within the trial cage populations, or there may have been some combination of metabolic enhancement and pathogen blocking (e.g. Caragata et al. 2013). Both the *w*Au and *w*Ri strains protect against artificially introduced infections of RNA viruses in *Drosophila* in laboratory settings (Osborne et al. 2009; Martinez et al. 2014), although there is little evidence that naturally-occurring *Wolbachia* infection protects against DNA viruses (Graham et al. 2012). Although we were able to detect some viruses from individual *D. simulans* sampled from our experimental cages these were only found at low frequency. The high microbe exposure condition applied some months prior to sampling from the population cages had no apparent impact on virus prevalence and there was no significant association with *Wolbachia* infection status, although resultant sample sizes were small.

The magnitude of the positive fitness effect associated with *Wolbachia* double infection in cage populations of *Ae. albopictus* reported by (Dobson et al. 2002) is exceptionally large. As infection frequency of *Wolbachia* in *Ae. albopictus* field populations has been found to be at, or very close to 100% (Kittayapong et al. 2002; Zouache et al. 2011), it is possible these mosquito hosts are reliant on *Wolbachia* (cf. Martinez et al. 2016) and that the relationship is tending towards an obligate rather than a facultative one. By contrast, field populations of *D. simulans* are polymorphic for *Wolbachia* infection around the world (Ballard 2004; Kriesner et al. 2013) and with the higher equilibrium frequency of the more recently spreading *w*Ri infection primarily explained by a balance between significant CI and maternal transmission loss. As the magnitude of fitness enhancement identified for the trial populations in this study is still very large, this underscores the potential utility of *D. simulans* as a useful model system to investigate conditional *Wolbachia* fitness effects and the context under which they arise. Additionally, an experimental system that yields *F* on the order of 1.19 or greater would likely only require census population sizes of around 1,000 adult individuals to be maintained in order to sufficiently mitigate stochastic effects.

Although we failed to detect a consistent difference in fitness effects between our high and low exposure cages, *Wolbachia* enhancement of host fitness might nevertheless still involve interactions with other microbes, whose identity and transmission mode remain undefined. Nora Virus was found to have little effect on host fecundity or longevity in *D. melanogaster* lab stocks, and was mainly present in the fly gut (Habayeb et al. 2009) so perhaps *Wolbachia* will be ineffective in blocking or slowing transmission of this virus. Drosophila C virus is detrimental to *D. melanogaster*, at least when artificially injected (Arnold et al. 2013) but was not detected in field populations of *D. simulans* (Webster et al. 2015). Little is currently known regarding effects of other pathogens on either of these *Drosophila* model species.

Given the potential public health benefits from applied use of *Wolbachia* in field populations of *Aedes aegypti* and possibly other human disease vectors, a sound understanding of the assets and liabilities that make up the 'balance sheet' of *Wolbachia* effects on their hosts is needed to give these programs the best chance of success. For example it is currently unclear whether a fundamental trade-off between strength of pathogen blocking and effectiveness of reproductive manipulation exists, or whether these are essentially independent phenomena (Hoffmann et al. 2015). Information about these *Wolbachia* effects is inherently difficult to obtain from the field because they may be conditional on factors such as temperature, availability of certain nutrients or host population structure. An experimental system where most of these factors can be controlled has considerable potential to isolate and identify those that are of fundamental importance.

Cage	Initial number of mated adult females			Initial infection frequency	Target microbe exposure level
	Uninfected	*w*Au-infected	*w*Ri-infected		
'A'	650	350	-	35%	'High'
'B'	1440	-	60	4%	'High'
'C'	650	350	-	35%	'High'
'D'	1440	-	60	4%	'High'
'E'	650	350	-	35%	'Low'
'F'	1440	-	60	4%	'Low'
'G'	650	350	-	35%	'Low'
'H'	1440	-	60	4%	'Low'

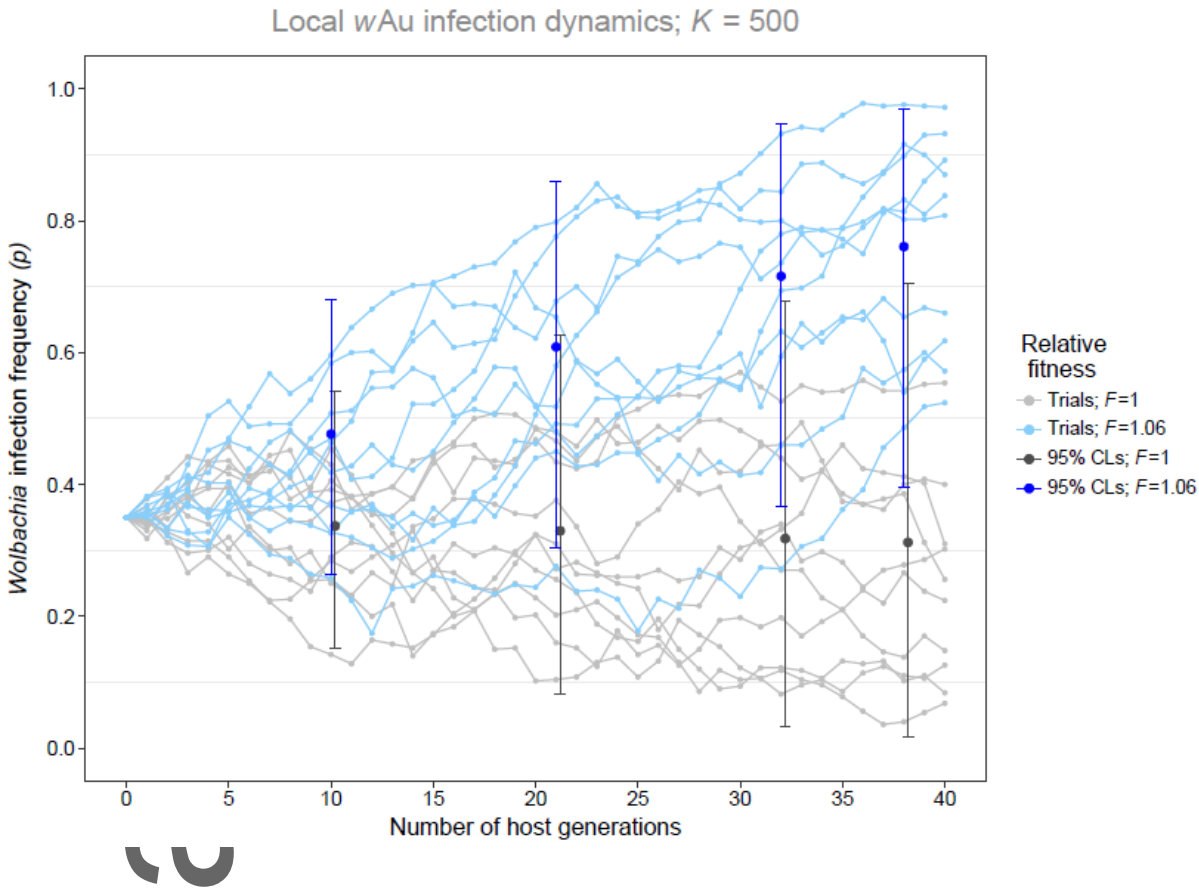
Table 1. Initial population composition of experiment cages

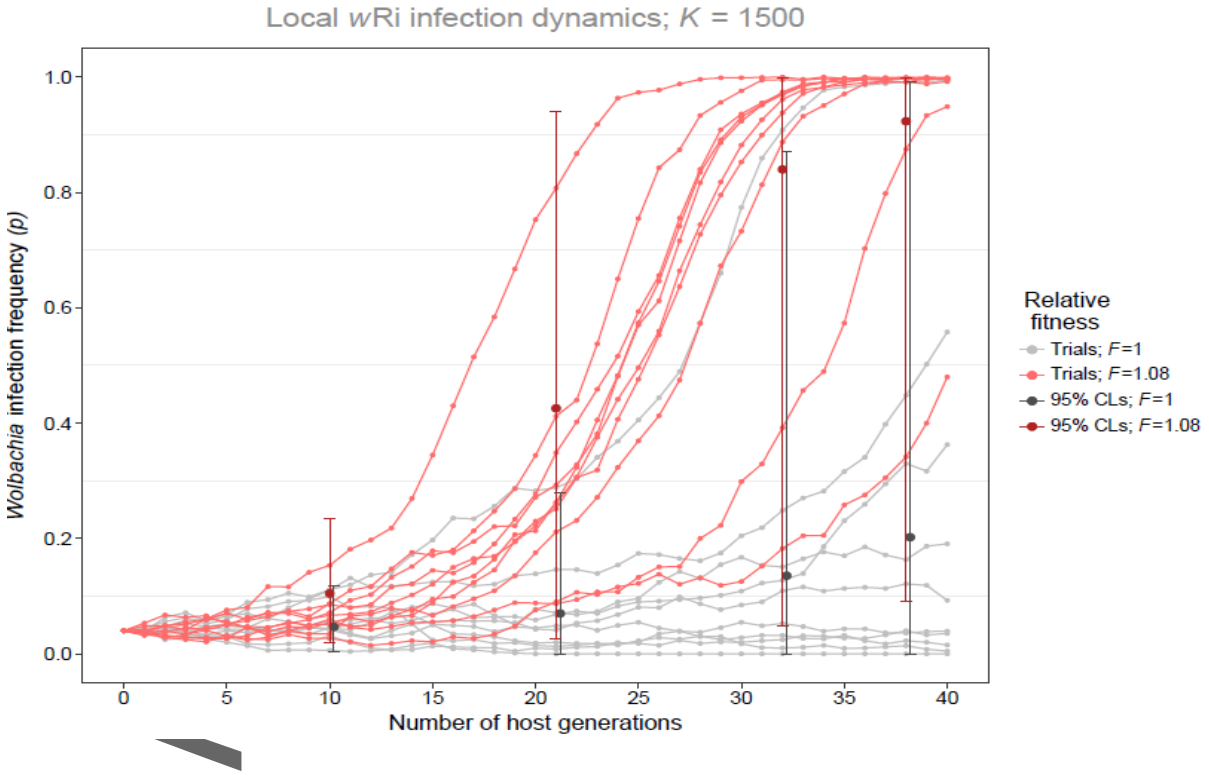
Cage	Estimates of effective population size (N_e)			
	Trial type	N_e [Jónás]	N_e [W]	N_e [JR]
'A'	*w*Au	4,336	15,571	1,007
'B'	*w*Ri	2,490	2,493	885
'C'	*w*Au	6,867	20,119	1,032

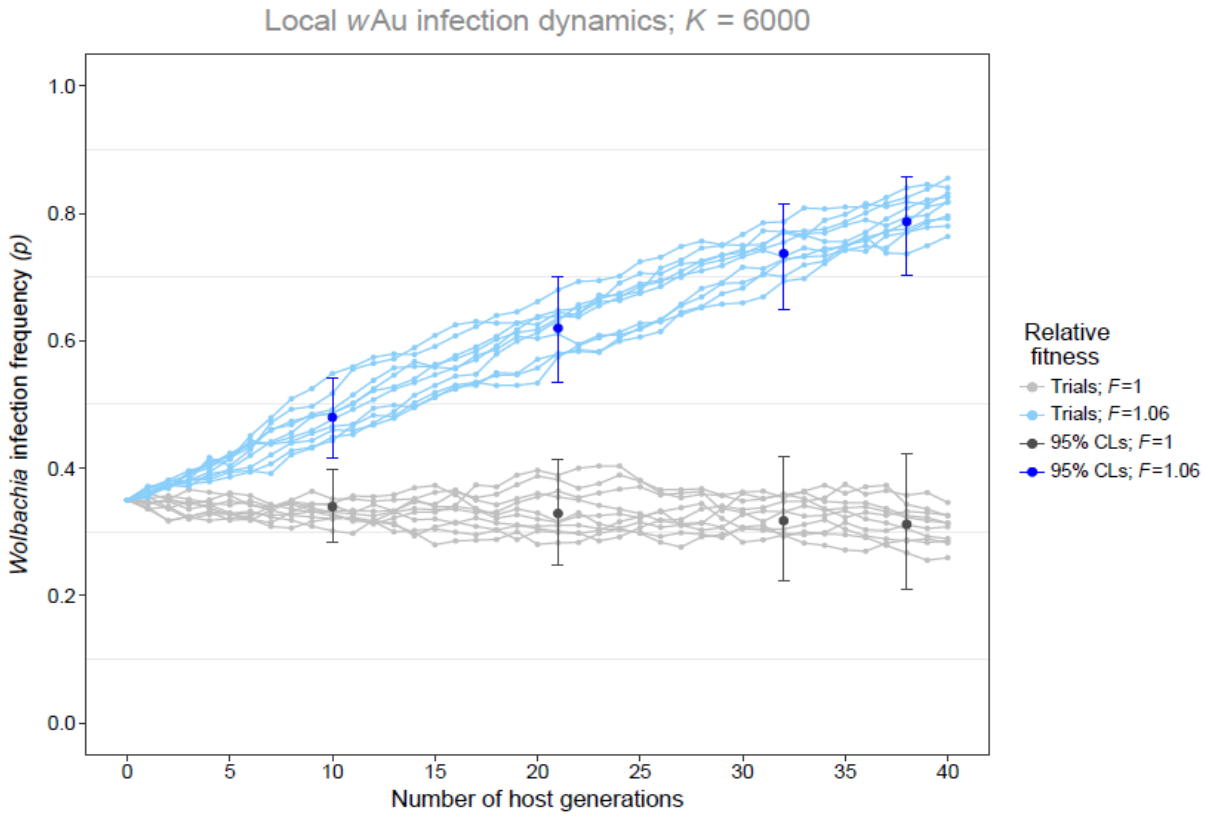
'D'	*w*Ri	1,431	2,027	684
'E'	*w*Au	2,848	5,442	879
'F'	*w*Ri	2,162	4,531	835
'G'	*w*Au	-174,676	68,761	1,305
'H'	*w*Ri	1,084	1,228	582

Table 2. Estimates for the effective population size of the 8 trial cage populations over approximately 7 months, estimated as equivalent to 22 generations using three alternative methods: N_e [Jónás] (Jónás et al. 2016), N_e [W] (Waples 1989), and N_e [JR] (Jorde and Ryman 2007). Only the N_e [Jónás] method specifically accounts for the additional variance component introduced by pooled sequencing.

Figure 1A-D. Stochastic model outcomes comparing effects of positive or neutral relative fitness (*F*) on infection dynamics over 40 host generations (10 random trials each) and mean average of Monte Carlo simulations (100,000 replicates) with associated 95% confidence limits (CLs) for F_{10}, F_{21}, F_{32} and F_{38} generations. For all simulations, μ = 0.003. (A) Small capacity environment: *w*Au dynamics with N_i = 500, p_i = 0.35, *H* = 100% (i.e. no CI effect), *F* = 1 or 1.06, *K* = 500. (B) Small capacity environment: *w*Ri dynamics with N_i = 1500, p_i = 0.04, *H* = 55%, *F* = 1 or 1.08, *K* = 1500. (C) Larger capacity environment: *w*Au dynamics with N_i = 1500, p_i = 0.35, *H* = 100%, *F* = 1 or 1.06, *K* = 6000. (D) Larger capacity environment: *w*Ri dynamics with N_i = 1500, p_i = 0.04, *H* = 55%, *F* = 1 or 1.08, *K* = 6000.







(D)

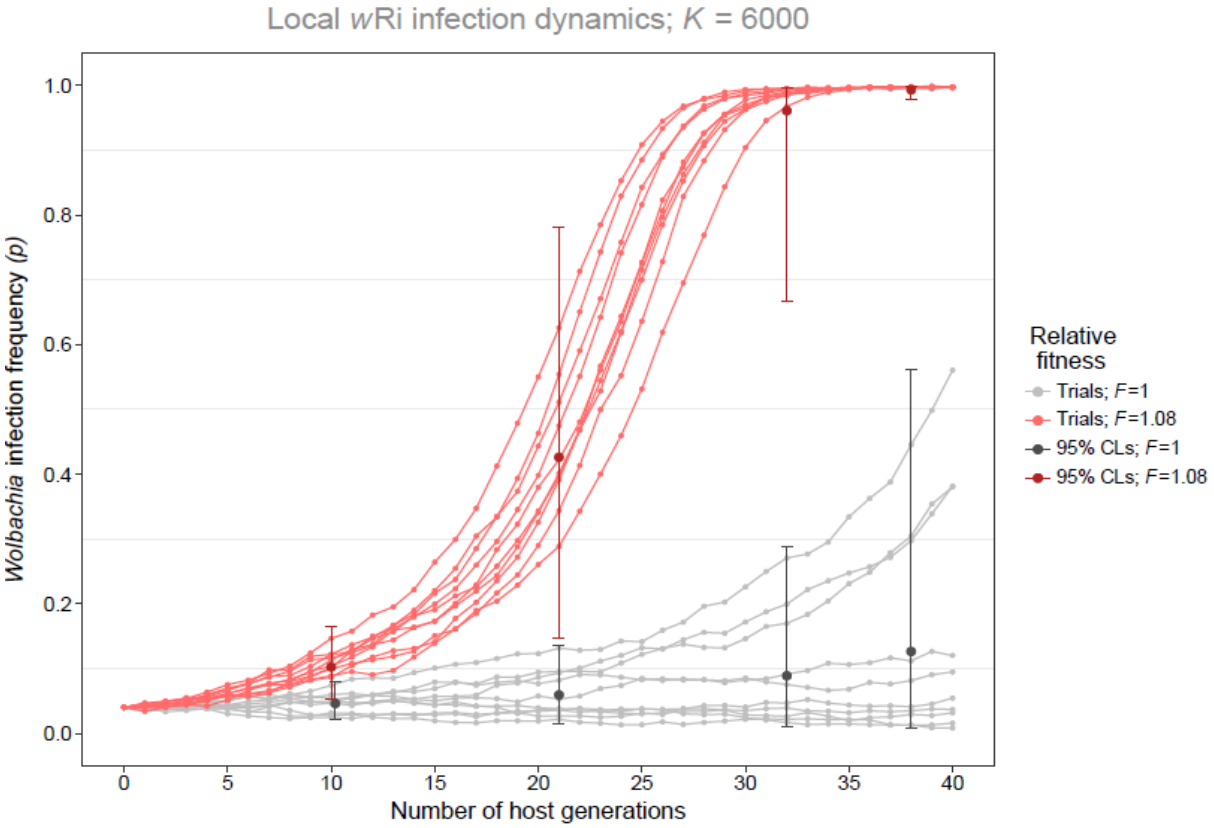
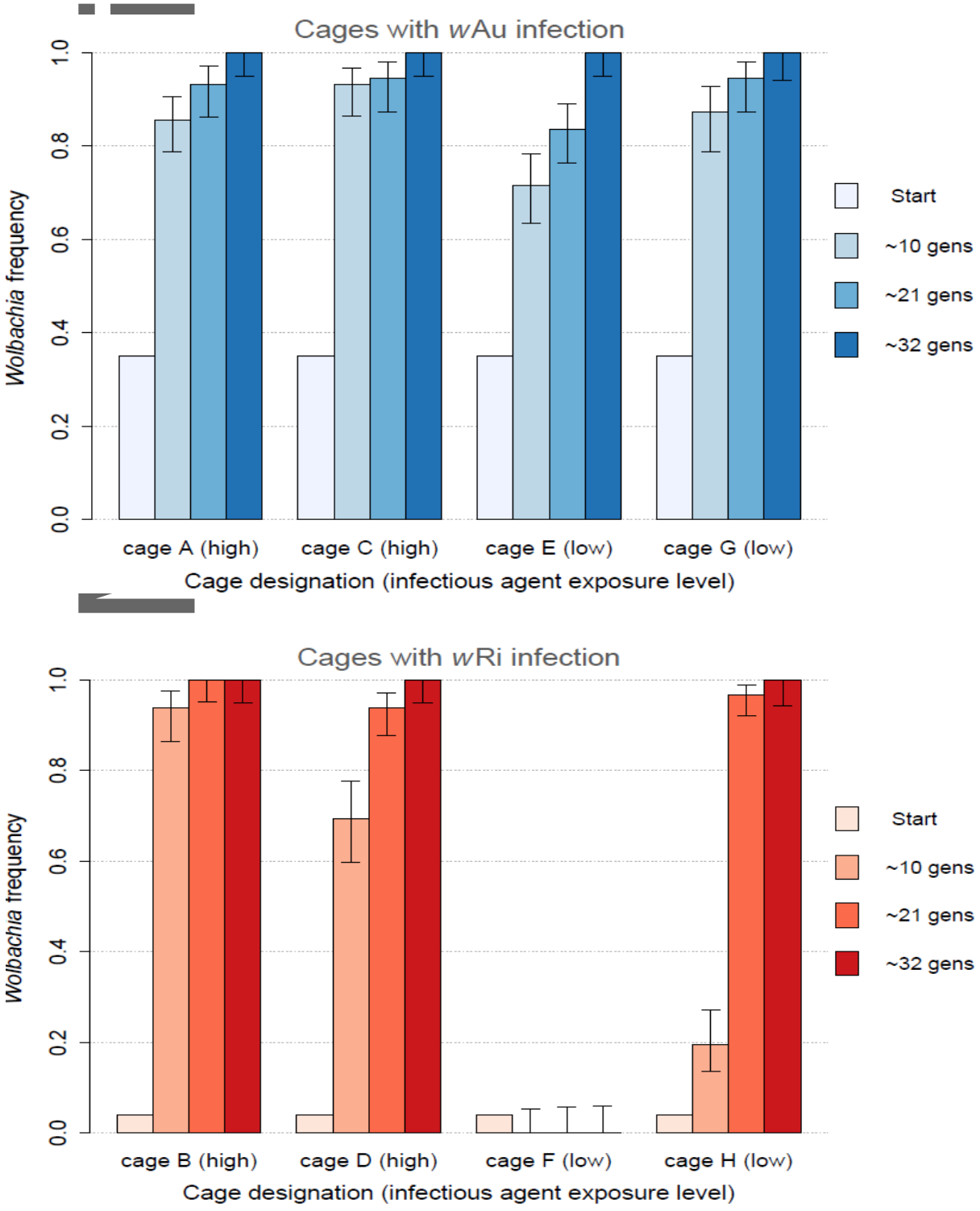


Figure 2. Infection frequencies of samples collected from experiment cages at approximately 3½ month intervals (April, July and November) estimated to coincide with 10, 21 and 32 host generations for cages with (A) *w*Au infection and (B) *w*Ri infection. Error bars indicate 95% confidence intervals.



References:

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Ant, T. H., C. S. Herd, V. Geoghegan, A. A. Hoffmann, and S. P. Sinkins. 2018. The *Wolbachia* strain *w*Au provides highly efficient virus transmission blocking in *Aedes aegypti*. PLOS Pathogens 14:e1006815.

Arnold, P. A., K. N. Johnson, and C. R. White. 2013. Physiological and metabolic consequences of viral infection in *Drosophila melanogaster*. Journal of Experimental Biology 216:3350-3357.

Ashburner, M., K. G. Golic, and R. S. Hawley. 2005. *Drosophila*: a laboratory handbook. Second edition. Cold Spring Harbor Laboratory Press, New York.

Axford, J. K., P. A. Ross, H. L. Yeap, A. G. Callahan, and A. A. Hoffmann. 2016. Fitness of *w*AlbB *Wolbachia* infection in *Aedes aegypti*: Parameter estimates in an outcrossed background and potential for population invasion. American Journal of Tropical Medicine and Hygiene 94:507-516.

Ballard, J. W. O. 2004. Sequential evolution of a symbiont inferred from the host: *Wolbachia* and *Drosophila simulans*. Molecular Biology and Evolution 21:428-442.

Barton, N. H. and M. Turelli. 2011. Spatial waves of advance with bistable dynamics: Cytoplasmic and genetic analogues of allee effects. American Naturalist 178:E48-E75.

Beckmann, J. F., J. A. Ronau, and M. Hochstrasser. 2017. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. Nature Microbiology 2:17007.

Brownlie, J. C., B. N. Cass, M. Riegler, J. J. Witsenburg, I. Iturbe-Ormaetxe, E. A. McGraw, and S. L. O'Neill. 2009. Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia pipientis*, during periods of nutritional stress. PLoS Pathogens 5:e1000368.

Caragata, E. P., E. Rancès, L. M. Hedges, A. W. Gofton, K. N. Johnson, S. L. O'Neill, and E. A. McGraw. 2013. Dietary cholesterol modulates pathogen blocking by *Wolbachia*. PLoS Pathogens 9:e1003459.

Carrington, L. B., J. R. Lipkowitz, A. A. Hoffmann, and M. Turelli. 2011. A re-examination of *Wolbachia*-induced cytoplasmic incompatibility in California *Drosophila simulans*. PLoS One 6(7):e22565.

Caspari, E. and G. S. Watson. 1959. On the evolutionary importance of cytoplasmic sterility in mosquitoes. Evolution 13:568-570.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124-3140.

Champion de Crespigny, F. E. C. and N. Wedell. 2007. Mate preferences in *Drosophila* infected with *Wolbachia*? Behavioral Ecology and Sociobiology 61:1229-1235.

Charlat, S., E. A. Hornett, J. H. Fullard, N. Davies, G. K. Roderick, N. Wedell, and G. D. D. Hurst. 2007. Extraordinary flux in sex ratio. Science 317:214.

Corbin, C., E. R. Heyworth, J. Ferrari, and G. D. D. Hurst. 2017. Heritable symbionts in a world of varying temperature. Heredity 118:10-20.

Dobson, S. L., E. J. Marsland, and W. Rattanadechakul. 2002. Mutualistic *Wolbachia* infection in *Aedes albopictus*: Accelerating cytoplasmic drive. Genetics 160:1087-1094.

Egas, M., F. Vala, and J. A. J. Breeuwer. 2002. On the evolution of cytoplasmic incompatibility in haplodiploid species. Evolution 56:1101-1109.

Fast, E. M., M. E. Toomey, K. Panaram, D. Desjardins, E. D. Kolaczyk, and H. M. Frydman. 2011. *Wolbachia* enhance *Drosophila* stem cell proliferation and target the germline stem cell niche. Science 334:990-992.

Fenton, A., K. N. Johnson, J. C. Brownlie, and G. D. D. Hurst. 2011. Solving the *Wolbachia* paradox: Modeling the tripartite interaction between host, *Wolbachia*, and a natural enemy. American Naturalist 178:333-342.

Fine, P. E. M. 1978. Dynamics of symbiote-dependent cytoplasmic incompatibility in culicine mosquitoes. Journal of Invertebrate Pathology 31:10-18.

Graham, R. I., D. Grzywacz, W. L. Mushobozi, and K. Wilson. 2012. *Wolbachia* in a major African crop pest increases susceptibility to viral disease rather than protects. Ecology Letters 15:993-1000.

Habayeb, M. S., R. Cantera, G. Casanova, J. O. Ekstrom, S. Albright, and D. Hultmark. 2009. The *Drosophila* Nora virus is an enteric virus, transmitted via feces. Journal of Invertebrate Pathology 101:29-33.

Hamm, C. A., D. J. Begun, A. Vo, C. C. R. Smith, P. Saelao, A. O. Shaver, J. Jaenike, and M. Turelli. 2014. *Wolbachia* do not live by reproductive manipulation alone: infection polymorphism in *Drosophila suzukii* and *D. subpulchrella*. Molecular Ecology 23:4871-4885.

Harcombe, W. and A. A. Hoffmann. 2004. *Wolbachia* effects in *Drosophila melanogaster*: in search of fitness benefits. Journal of Invertebrate Pathology 87:45-50.

Herre, E. A., N. Knowlton, U. G. Mueller, and S. A. Rehner. 1999. The evolution of mutualisms: exploring the paths between conflict and cooperation. Trends in Ecology & Evolution 14:49-53.

Himler, A. G., T. Adachi-Hagimori, J. E. Bergen, A. Kozuch, S. E. Kelly, B. E. Tabashnik, E. Chiel, V. E. Duckworth, T. J. Dennehy, E. Zchori-Fein, et al. 2011. Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. Science 332:254-256.

Hoffmann, A. A., D. Clancy, and J. Duncan. 1996. Naturally-occurring *Wolbachia* infection in *Drosophila simulans* that does not cause cytoplasmic incompatibility. Heredity 76:1-8.

Hoffmann, A. A., M. Hercus, and H. Dagher. 1998. Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. Genetics 148:221-231.

Hoffmann, A. A., I. Iturbe-Ormaetxe, A. G. Callahan, B. L. Phillips, K. Billington, J. K. Axford, B. Montgomery, A. P. Turley, and S. L. O'Neill. 2014. Stability of the *w*Mel *Wolbachia* infection following invasion into *Aedes aegypti* populations. PLoS Neglected Tropical Diseases 8:e3115.

Hoffmann, A. A., B. L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P. H. Johnson, F. Muzzi, M. Greenfield, M. Durkan, Y. S. Leong, Y. Dong, et al. 2011. Successful

establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. Nature 476:454-457.

Hoffmann, A. A., P. A. Ross, and G. Rašić. 2015. *Wolbachia* strains for disease control: ecological and evolutionary considerations. Evolutionary Applications 8:751-768.

Hoffmann, A. A. and M. Turelli. 1997. Cytoplasmic incompatibility in insects. Pp. 42-80 *in* S. L. O'Neill, A. A. Hoffmann, and J. H. Werren, eds. Influential Passengers: Inherited Microorganisms and Arthropod Reproduction. Oxford University Press, New York.

Hoffmann, A. A., M. Turelli, and L. G. Harshman. 1990. Factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. Genetics 126:933-948.

Hood, G. M. (2011) PopTools version 3.2.5. Available on the internet. URL http://www.poptools.org.

Hornett, E. A., S. Charlat, A. M. R. Duplouy, N. Davies, G. K. Roderick, N. Wedell, and G. D. D. Hurst. 2006. Evolution of male-killer suppression in a natural population. PLoS Biology 4:1643-1648.

Hosokawa, T., R. Koga, Y. Kikuchi, X. Y. Meng, and T. Fukatsu. 2010. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. Proceedings of the National Academy of Sciences of the United States of America 107:769-774.

Hui, C. 2006. Carrying capacity, population equilibrium, and environment's maximal load. Ecological Modelling 192:317-320.

Jaenike, J. 2012. Population genetics of beneficial heritable symbionts. Trends in Ecology & Evolution 27:226-232.

Jansen, V. A. A., M. Turelli, and H. C. J. Godfray. 2008. Stochastic spread of *Wolbachia*. Proceedings of the Royal Society B-Biological Sciences 275:2769-2776.

Jónás, Á., T. Taus, C. Kosiol, C. Schlötterer, and A. Futschik. 2016. Estimating the effective population size from temporal allele frequency changes in experimental evolution. Genetics 204:723-735.

Jorde, P. E. and N. Ryman. 2007. Unbiased estimator for genetic drift and effective population size. Genetics 177:927-935.

Joshi, A., M. H. Do, and L. D. Mueller. 1999. Poisson distribution of male mating success in laboratory populations of *Drosophila melanogaster*. Genetical Research 73:239-249.

Kittayapong, P., V. Baimai, and S. L. O'Neill. 2002. Field prevalence of *Wolbachia* in the mosquito vector *Aedes albopictus*. American Journal of Tropical Medicine and Hygiene 66:108-111.

Klasson, L., J. Westberg, P. Sapountzis, K. Nasiund, Y. Lutnaes, A. C. Darby, Z. Veneti, L. M. Chen, H. R. Braig, R. Garrett, et al. 2009. The mosaic genome structure of the *Wolbachia w*Ri strain infecting *Drosophila simulans*. Proceedings of the National Academy of Sciences of the United States of America 106:5725-5730.

Kofler, R., R. V. Pandey, and C. Schlotterer. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435-3436.

Kriesner, P., A. A. Hoffmann, S. F. Lee, M. Turelli, and A. R. Weeks. 2013. Rapid sequential spread of two *Wolbachia* variants in *Drosophila simulans*. PLoS Pathogens 9:e1003607.

Langmead, B. and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357-U354.

Lee, S. F., V. L. White, A. R. Weeks, A. A. Hoffmann, and N. M. Endersby. 2012. High-throughput PCR assays to monitor *Wolbachia* infection in the dengue mosquito (*Aedes aegypti*) and *Drosophila simulans*. Applied and Environmental Microbiology 78:4740-4743.

LePage, D. P., J. A. Metcalf, S. R. Bordenstein, J. M. On, J. I. Perlmutter, J. D. Shropshire, E. M. Layton, L. J. Funkhouser-Jones, J. F. Beckmann, and S. R. Bordenstein. 2017. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. Nature 543:243-247.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and P. Genome Project Data. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.

Markow, T. A., S. Beall, and L. M. Matzkin. 2009. Egg size, embryonic development time and ovoviviparity in *Drosophila* species. Journal of Evolutionary Biology 22:430-434.

Martinez, J., R. Cogni, C. Cao, S. Smith, C. J. R. Illingworth, and F. M. Jiggins. 2016. Addicted? Reduced host resistance in populations with defensive symbionts. Proceedings of the Royal Society B-Biological Sciences 283:10.

Martinez, J., B. Longdon, S. Bauer, Y.-S. Chan, W. J. Miller, K. Bourtzis, L. Teixeira, and F. M. Jiggins. 2014. Symbionts commonly provide broad spectrum resistance to viruses in insects: A comparative analysis of *Wolbachia* strains. PLoS Pathogens 10:e1004369.

Moriyama, M., N. Nikoh, T. Hosokawa, and T. Fukatsu. 2015. Riboflavin provisioning underlies *Wolbachia's* fitness contribution to its insect host. Mbio 6:1-8.

Nei, M. and F. Tajima. 1981. Genetic drift and estimation of effective population-size. Genetics 98:625-640.

Osborne, S. E., Y. S. Leong, S. L. O'Neill, and K. N. Johnson. 2009. Variation in antiviral protection mediated by different *Wolbachia* strains in *Drosophila simulans*. PLoS Pathogens 5:e1000656.

Peel, D., R. S. Waples, G. M. Macbeth, C. Do, and J. R. Ovenden. 2013. Accounting for missing data in the estimation of contemporary genetic effective population size (N-e). Molecular Ecology Resources 13:243-253.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS One 7:e37135.

Poinsot, D., K. Bourtzis, G. Markakis, C. Savakis, and H. Mercot. 1998. *Wolbachia* transfer from *Drosophila melanogaster* into *D. simulans*: Host effect and cytoplasmic incompatibility relationships. Genetics 150:227-237.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

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

Reynolds, K. T. and A. A. Hoffmann. 2002. Male age, host effects and the weak expression or nonexpression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. Genetical Research 80:79-87.

Ross, P. A., I. Wiwatanaratanabutr, J. K. Axford, V. L. White, N. M. Endersby-Harshman, and A. A. Hoffmann. 2017. Wolbachia infections in *Aedes aegypti* differ markedly in their response to cyclical heat stress. PLoS Pathogens 13:e1006006.

Schlötterer, C., R. Tobler, R. Kofler, and V. Nolte. 2014. Sequencing pools of individuals-mining genome-wide polymorphism data without big funding. Nature Reviews Genetics 15:749-763.

Sutton, E., S. Harris, J. Parkhill, and S. Sinkins. 2014. Comparative genome analysis of *Wolbachia* strain *w*Au. BMC Genomics 15:928.

Teixeira, L., A. Ferreira, and M. Ashburner. 2008. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. PLoS Biology 6(12):e1000002.

Turelli, M. 1994. Evolution of incompatibilty-inducing microbes and their hosts. Evolution 48:1500-1513.

Turelli, M. 2010. Cytoplasmic incompatibility in populations with overlapping generations. Evolution 64:232-241.

Turelli, M. and A. A. Hoffmann. 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature 353:440-442.

Turelli, M. and A. A. Hoffmann. 1995. Cytoplasmic incompatibility in *Drosophila simulans* - dynamics and parameter estimates from natural populations. Genetics 140:1319-1338.

Vala, F., M. Egas, J. A. J. Breeuwer, and M. W. Sabelis. 2004. *Wolbachia* affects oviposition and mating behaviour of its spider mite host. Journal of Evolutionary Biology 17:692-700.

Walker, T., P. H. Johnson, L. A. Moreira, I. Iturbe-Ormaetxe, F. D. Frentiu, C. J. McMeniman, Y. S. Leong, Y. Dong, J. Axford, P. Kriesner, et al. 2011. The *w*Mel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. Nature 476:450-453.

Waples, R. S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics 121:379-391.

Webster, C. L., F. M. Waldron, S. Robertson, D. Crowson, G. Ferrari, J. F. Quintana, J.-M. Brouqui, E. H. Bayne, B. Longdon, A. H. Buck, et al. 2015. The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. PLoS Biology 13:e1002210.

Weeks, A. R., M. Turelli, W. R. Harcombe, K. T. Reynolds, and A. A. Hoffmann. 2007. From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*. PLoS Biology 5(5):e114.

Woods, R. E., C. M. Sgro, M. J. Hercus, and A. A. Hoffmann. 2002. Fluctuating asymmetry, fecundity and development time in *Drosophila*: is there an association under optimal and stress conditions? Journal of Evolutionary Biology 15:146-157.

Wu, M., L. V. Sun, J. Vamathevan, M. Riegler, R. Deboy, J. C. Brownlie, E. A. McGraw, W. Martin, C. Esser, N. Ahmadinejad, et al. 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis w*Mel: A streamlined genome overrun by mobile genetic elements. PLoS Biology 2:327-341.

Zhang, H., K. J. Zhang, and X. Y. Hong. 2010. Population dynamics of noncytoplasmic incompatibility-inducing *Wolbachia* in *Nilaparvata lugens* and its effects on host adult life span and female fitness. Environmental Entomology 39:1801-1809.

Zhao, D. X., D. S. Chen, C. Ge, T. Gotoh, and X. Y. Hong. 2013. Multiple infections with *Cardinium* and two strains of *Wolbachia* in the spider mite *Tetranychus phaselus* Ehara: Revealing new forces driving the spread of *Wolbachia*. PLoS One 8(1) e54964:1-9.

Zouache, K., F. N. Raharimalala, V. Raquin, V. Tran-Van, L. H. R. Raveloson, P. Ravelonandro, and P. Mavingui. 2011. Bacterial diversity of field-caught mosquitoes, *Aedes albopictus* and *Aedes aegypti*, from different geographic regions of Madagascar. Fems Microbiology Ecology 75:377-389.