SUPPLEMENTAL APPENDIX 1

Criteria used:

1. Final infection frequency is higher than the initial infection frequency.
2. Two-thirds of the recorded infection frequency is above the initial infection frequency.

A function introduces sample size stochasticity to the model. The model is based on a discrete generation time (no overlapping generations), which is applicable to laboratory colony cages. Two levels of stochasticity were introduced based on:

1. The male (uninfected/infected) mated with a female.
2. Random sampling of offspring to set up the next generation.

Let N be the population size (male + female) and unstable equilibrium computed with no maternal transmission leakage and no viable offspring from incompatible crosses (dynamics driven only by fecundity/fitness cost).

An unstable of equilibrium of 0.15 was suggested by data from Xi and others[14], with no maternal transmission leakage, perfect cytoplasmic incompatibility, and fecundity/fitness cost of 15%. Likelihood was determined via 100,000 repetitions of the model with population size and mating stochasticity, and no enforcing of a constant sex ratio (no sexing). Forcing constant sex ratio does not really improve the likelihood much (not shown here) for the population sizes considered. The likelihood of detection (see Supplemental Table 1) can be observed to be dependent on several factors.

1. Sample size.
   a) If above unstable equilibrium, larger sample size will increase detection rate (power, reducing false negative).
   b) If below unstable equilibrium, larger sample size will decrease detection rate (reducing false positive).
2. Positively dependent on the starting infection frequency.
3. The true unstable equilibrium and the parameters giving rise to it.
4. Number of generations. Some simulated replicates required more time to follow expectations based on the invasion criteria, and/or some that appear to invade then crash.

The likelihood of observing invasion after 10 generations under different conditions is given in Supplemental Table 2.

SUPPLEMENTAL APPENDIX 2

We define β_{am} as the relative fitness of infected to uninfected males. We define an effective infection frequency that is an adjusted infection frequency taking into account the mating isolating effects of assortative mating and sperm competition.

Under the effects of assortative mating, if p_{am} is the effective infection frequency due to assortative mating, while p_I is the observed infection frequency, then

$$p_{am} = \frac{\beta_{am} p_I}{\beta_{am} p_I + (1 - p_I)}$$

This can be thought of as a weighted average method, with the infected males being reduced or increased by the β_{am} parameter. In the assortative mating experiment, we only considered uninfected females. The average observed hatch rate among uninfected female, H_{am}, is

$$H_{am} = (1 - p_{am} + p_{am} H) \times h$$

where H = average hatch rate of incompatible cross (uninfected females to infected males), and h = average absolute hatch rate of compatible cross (assume not affected by *Wolbachia*).

$1 - p_{am}$ is simply the hatch rate due to matings with uninfected males, while $p_{am}H$ is the additional observation due to incomplete cytoplasmic incompatibility. We assumed that the overall compatible hatch rate is not perfect, $h < 1$, hence the extra multiplier to restrict the model.

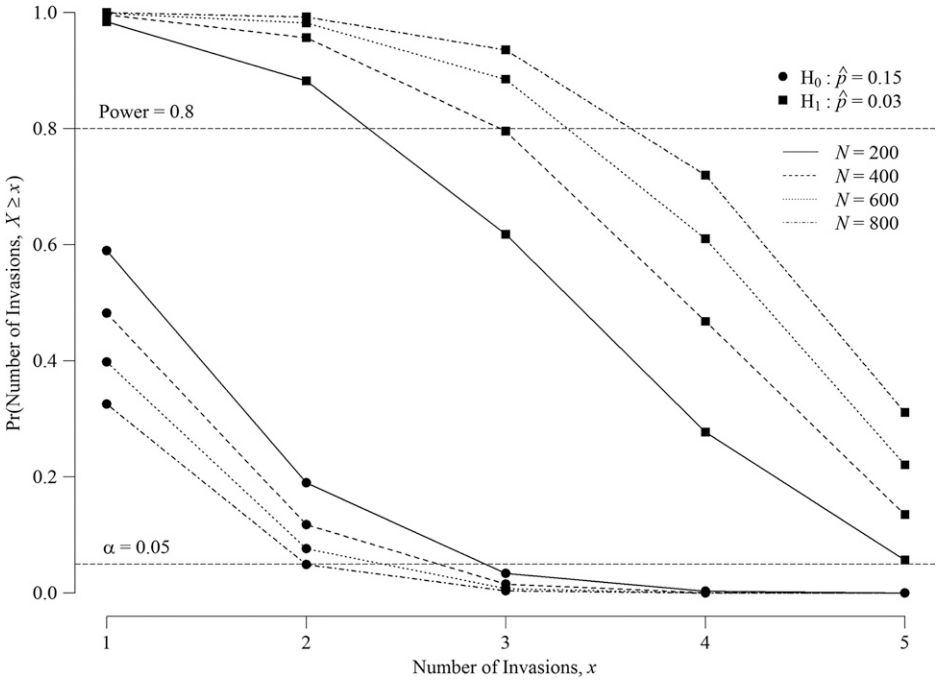
Thus, we can substitute p_{am}:

$$H_{am} = \frac{(1 - p_I + \beta_{am} p_I H) \times h}{\beta_{am} p_I + (1 + p_I)}$$

We can then formulate a linear equation by shifting the terms around:

$$\left(\frac{H_{am}}{h} - 1\right)(1 - p_I) = \beta_{am} \times \left(p_I H - p_I \frac{H_{am}}{h}\right) + 0$$

with the LHS of the equation as the Y-term and the multiplier to β_{am} as the X-term. This model has a Y-intercept of zero.



SUPPLEMENTAL FIGURE 1. Probability of observing $X \geq x$ invasions under H_0: $\hat{p} = 0.15$ (circles) or H_1: $\hat{p} = 0.03$ (squares) with an initial invasion frequency of $p_0 = 0.1$ across population sizes of N: 200 (solid line), 400 (dashed line), 600 (dotted line) and 800 (dot-dash line). Naturally, the probability of X being greater than or equal to x increases under H_1 or decreases under H_0 with increasing values of N. We chose $N = 400$, the lowest sample size with adequate power, 80% and Type I error, $\alpha < 0.05$ (see Supplemental Table 2 for simulated probability values).

SUPPLEMENTAL TABLE 1

Probability of seeing an invasion in a single replicate population of constant population size, N with varying unstable equilibrium

N	p_{invasion} (if $\hat{p} = 0.15$)	p_{invasion} (if $\hat{p} = 0.03$)
200	0.1634	0.5636
400	0.1233	0.6702
600	0.0965	0.7393
800	0.0757	0.7915

Starting infection frequency, $p_0 = 0.1$.

SUPPLEMENTAL TABLE 2

Likelihood of observing invasion after 10 generations at starting infection frequencies, $p_0 = 0.1$ and population sizes, N

	$\Pr(X \geq x)$ if $\hat{p} = 0.15$					$\Pr(X \geq x)$ if $\hat{p} = 0.03$				
N	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$
200	0.5901	0.1899	**0.0336**	**0.0031**	**0.0001**	*0.9842*	*0.8819*	0.6179	0.2770	0.0569
400	0.4821	0.1179	**0.0155**	**0.0010**	**0.0000**	*0.9961*	*0.9564*	*0.7953*	0.4679	0.1352
600	0.3980	0.0764	**0.0077**	**0.0004**	**0.0000**	*0.9988*	*0.9817*	*0.8849*	0.6102	0.2208
800	0.3255	**0.0492**	**0.0039**	**0.0002**	**0.0000**	*0.9996*	*0.9921*	*0.9354*	0.7198	0.3107

X = number of invasions observed in five replicate populations of size N each generation and starting *Wolbachia* infection frequency, $p_0 = 0.1$. Values in bold represent the likelihood of false positives that are below 5% (based on H_0: $\hat{p} = 0.15$). For example, for $N = 400$, there is a 1.55% chance of observing at least three invasions in five replicated population of 400 individuals with equal sex ratio and initial *Wolbachia* infection frequency, $p_0 = 0.1$, if the unstable equilibrium is truly 0.15 (due to fecundity/fitness cost of 15%). Values in italic represent the power of detection around 80% and above (based on H_1: $\hat{p} = 0.03$). Probabilities are calculated using binomial sampling based on single trial probabilities in Supplemental Table 1. The unstable equilibrium, $\hat{p}$ at 0.15 was based on Xi and others[14], while $\hat{p} = 0.03$ was based on our data collected during the "Quiescent egg viability" experiment.

SUPPLEMENTAL TABLE 3

Cox regressions for "adult survival in groups" for females sorted by line then block

Line	Block*	β†	e^{β}‡	e^{β} lower 95% CI	e^{β} upper 95% CI	z	Pr(>|z|)
*w*AlbB	1	0.341 ± 0.417	1.41	0.62	3.18	0.82	0.414
	2	0.773 ± 0.416	2.17	0.96	4.89	1.86	0.063
	3	2.145 ± 0.162	8.54	6.22	11.73	13.27	< 0.001
*w*Mel	1	0.688 ± 0.401	1.99	0.91	4.37	1.71	0.087
	2	1.328 ± 0.398	3.77	1.73	8.23	3.34	< 0.001
	3	0.969 ± 0.153	2.64	1.95	3.56	6.34	< 0.001
*w*MelPop	1	0.444 ± 0.441	1.56	0.66	3.70	1.01	0.314
	2	3.074 ± 0.371	21.63	10.45	44.75	8.29	< 0.001
	3	3.085 ± 0.240	21.87	13.66	35.01	12.85	< 0.001

95% CI = 95% confidence interval.
*Block 1 = 0–20 days, block 2 = 21–40 days, block 3 = > 40 days.
†Survival curve slope ± standard error.
‡Hazard ratio.

SUPPLEMENTAL TABLE 4

Pairwise comparisons within blocks of time for females in the "adult survival in groups" experiment using general linear hypothesis tests

Block	Null hypothesis	Estimate	SE	z	Pr(>|z|)*
1 (0–20 days)	*w*AlbB − CNS = 0	0.341	0.417	0.82	0.845
	*w*Mel − CNS = 0	0.688	0.401	1.71	0.314
	*w*MelPop − CNS = 0	0.444	0.441	1.01	0.743
	*w*Mel − *w*AlbB = 0	0.348	0.335	1.04	0.726
	*w*MelPop − *w*AlbB = 0	0.104	0.382	0.27	0.993
	*w*MelPop − *w*Mel = 0	−0.244	0.365	−0.67	0.909
2 (21–40 days)	*w*AlbB − CNS = 0	0.773	0.416	1.86	0.233
	*w*Mel − CNS = 0	1.328	0.398	3.34	**0.004**
	*w*MelPop − CNS = 0	3.074	0.371	8.29	**< 0.001**
	*w*Mel − *w*AlbB = 0	0.555	0.285	1.95	0.197
	*w*MelPop − *w*AlbB = 0	2.301	0.245	9.38	**< 0.001**
	*w*MelPop − *w*Mel = 0	1.746	0.214	8.16	**< 0.001**
3 (> 40 days)	*w*AlbB − CNS = 0	2.145	0.162	13.27	**< 0.001**
	*w*Mel − CNS = 0	0.969	0.153	6.34	**< 0.001**
	*w*MelPop − CNS = 0	3.085	0.240	12.85	**< 0.001**
	*w*Mel − *w*AlbB = 0	−1.176	0.137	−8.56	**< 0.001**
	*w*MelPop − *w*AlbB = 0	0.940	0.202	4.65	**< 0.001**
	*w*MelPop − *w*Mel = 0	2.116	0.224	9.47	**< 0.001**

SE = standard error.
*Significant z values are shown in bold.