Positive fitness effects help explain the broad range of *Wolbachia* prevalences in natural populations

P. Karisto[1,†], A. Duplouy[2], C. de Vries[1,‡], H. Kokko[1,3]

Article

[1] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.
[2] Insect Symbiosis Ecology and Evolution, Organismal and Evolutionary Biology Research Program, University of Helsinki, Helsinki, Finland.
[3] Konrad Lorenz Institute of Ethology, University of Veterinary Medicine, Vienna, Austria.
† Current address: Plant Health, Natural Resources Institute Finland, Jokioinen, Finland.
‡ Current address: Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland.

Corresponding author: P. Karisto
Email address: petteri.karisto@luke.fi

Keywords: Cytoplasmic incompatibility, haplodiploid, diplodiploid

Running title: Low frequency *Wolbachia*

Abstract

The bacterial endosymbiont *Wolbachia* is best known for its ability to modify its host's reproduction by inducing cytoplasmic incompatibility (CI) to facilitate its own spread. Classical models predict either near-fixation of costly *Wolbachia* once the symbiont has overcome a threshold frequency (invasion barrier), or *Wolbachia* extinction if the barrier is not overcome. However, natural populations do not all follow this pattern: *Wolbachia* can also be found at low frequencies that appear stable over time. *Wolbachia* is known to have pleiotropic fitness effects (beyond CI) on its hosts. Existing models typically focus on the possibility that these are negative. Here we consider the possibility that the symbiont provides direct benefits to infected females (e.g. resistance to pathogens) in addition to CI. We discuss an underappreciated feature of *Wolbachia* dynamics: that CI with pleiotropy can produce low-frequency (<0.5) stable equilibria, when pleiotropy allows infected hosts to enjoy some fitness advantages. Additionally, without direct positive fitness effect, any stable equilibrium close to 0.5 will be sensitive to perturbations, which make such equilibria unlikely to sustain in nature. The results hold for both diplodiploid and different haplodiploid versions of CI. We suggest that insect populations showing low-frequency *Wolbachia* infection rates might host CI-inducing symbiotic strains providing additional (hidden or known) benefits to their hosts, especially when classical explanations (ongoing invasion, source-sink dynamics) have been ruled out.

Introduction

Maternally inherited endosymbiontic bacteria of genus *Wolbachia* are widespread in arthropods and nematodes (Hilgenboecker et al., 2008; Zug and Hammerstein, 2012; Weinert et al., 2015). Since the discovery of *Wolbachia*-induced cytoplasmic incompatibility (CI, Laven, 1956; Yen and Barr, 1971) – embryonic mortality in the offspring of parents of different infection status – the symbionts have attained considerable empirical as well as theoretical interest. CI is an obvious and dramatic fitness effect, but *Wolbachia*'s effects are often pleiotropic, impacting the fitness of its hosts directly and regardless of the type of mates that the host reproduces with (for example Hoffmann et al., 1990). Assuming that pleiotropy includes direct costs of infection, it becomes at first sight difficult to understand how *Wolbachia* can spread and reach high frequencies. However, theory soon showed that costly *Wolbachia* infections show positive frequency-dependence. At low frequencies, costs predominate and *Wolbachia* cannot spread, but if the initial frequency of *Wolbachia* is sufficiently high (above an invasion barrier), then the frequency-dependent success favours the CI-inducing symbiont sufficiently for it to spread (Fine, 1978; Caspari and Watson, 1959). At equilibrium it will be close to fixation, only prevented from fixing by some infections failing to be passed on to offspring (Hoffmann et al., 1990; Engelstädter and Telschow, 2009). This fits in with the general expectation that positive frequency dependence makes it easy to explain either absence or (near) fixation of a trait (Lehtonen and Kokko, 2012).

In many empirical systems, *Wolbachia* are indeed very prevalent in their host populations (Hoffmann et al., 1990; Jeong et al., 2009; Duplouy and Brattström, 2018; Deng et al., 2021) in line with existing theory. However, the recent accumulation of studies on the prevalence and penetrance of *Wolbachia* in diverse host species and populations show a wider range of infection prevalences. Sazama et al. (2019) have reported that infection prevalence of *Wolbachia* is likely to remain below 0.5 in the majority of insect species

(Tagami and Miura, 2004; Sun et al., 2007; Arthofer et al., 2009; Hughes et al., 2011), including systems where the spread of the *Wolbachia* is potentially on-going (Duplouy et al., 2010). There are also examples of low prevalence remaining reasonably stable over time (Duplouy et al., 2015). Although some cases of low and apparently stable prevalence might be explicable as a result of migration to sink populations from sources where symbiont prevalence is high (Flor et al., 2007; Telschow et al., 2007), it appears difficult to consider spatial dynamics a sufficient explanation for widespread cases of empirically documented low prevalences.

Here we show that spatial heterogeneity is not needed for models to produce low stable frequencies of CI-inducing symbionts, once one allows the effects of infection on the host to be positive. Studies allowing positive fitness effect of CI inducing symbiont are rare. However, Zug and Hammerstein (2018) investigated the effect of positive fitness on the invasion dynamics of different reproductive parasites into virgin populations and against each other. They demonstrated the key role of relative fitness in determining invasion conditions of the parasites. The effect is often observed as lower (or absent) invasion threshold frequency. While their focus was mostly on the invasibility of the parasites, we aim to widen the understanding of the resulting stable equilibrium frequencies when positive fitness effects are present.

We first revisit classic CI models (following the notation of Engelstädter and Telschow, 2009) and show that the assumption structure, where infected hosts experience either no or negative fitness effects, cannot equilibrate at infection prevalences below 0.5 (as stated by Turelli, 1994, for a diplodiploid case). We also argue that within the range predicted by the classic model (0.5 to 1), the lower end of the range is difficult to maintain in the presence of stochastic fluctuations in infection prevalence, as the invasion barrier will in these cases be close to the stable equilibrium. We then show that when we allow infected hosts to benefit from their infection status, the stable equilibria can exist below

0.5 in wide range of cases, including situations with and without the invasion barrier. The results of the classical diplodiploid model of CI are quite similar to different haplodiploid models (with an exception that we discuss). Our observations of low frequency stable equilibria are not entirely novel: the analysis of Zug and Hammerstein (2018) includes a demonstration of such a situation in diplodiploid system (their Supplementary Fig. S1), albeit without any written comment as the focus of these authors was elsewhere. Hence, our aim is to strengthen the appreciation of the dynamic consequences of positive fitness effects in the dynamics of CI-inducing symbionts.

Models and analysis

Diplodiploid model

Hoffmann et al. (1990) presented a seminal model for the spread of CI-inducing symbionts in diploid-diploid species, building upon a slightly different model of Fine (1978). We use the discrete time model of Hoffmann et al. (1990) but update it to the notation of Engelstädter and Telschow (2009). *Wolbachia*-bearing females have relative fecundity f compared to the uninfected baseline (which we set to unity). *Wolbachia* is transmitted maternally to a proportion t of eggs. Under rules of CI, an uninfected egg that is fertilized by a *Wolbachia*-modified sperm (produced by an infected male) dies with probability L, while *Wolbachia*-infected eggs survive irrespective of the status of the fertilizing sperm (to be precise, this requires sperm to be either from a *Wolbachia*-free father or from a father with the same or a compatible *Wolbachia* strain as the egg; for simplicity we follow earlier studies and consider only one strain in our model).

Our model assumes equal frequency (p_T) of *Wolbachia* in males and females in the current generation (at time T). We track four types of matings (the mother and the father can each be uninfected or infected) and their outcome in terms of infected and un-

infected offspring production. Only infected mothers can produce infected offspring, and the total of such offspring is $p_T ft$ (sperm genotype does not matter for these mothers). The denominator consists of all offspring from all matings:

$$\begin{aligned} p_{T+1} &= \frac{p_T ft}{\underbrace{(1-p_T)^2}_{\text{UF x UM}} + \underbrace{(1-p_T)p_T(1-L)}_{\text{UF x IM}} + \underbrace{p_T(1-p_T)f}_{\text{IF x UM}} + \underbrace{p_T^2 f[t+(1-t)(1-L)]}_{\text{IF x IM}}} \\ &= \frac{p_T ft}{p_T f(1-(1-t)p_T L) + (1-p_T)(1-p_T L)}. \end{aligned} \tag{1}$$

The equation (1) has three equilibria:

$$\hat{p}_0 = 0,\ \hat{p}_1 = \frac{-B-\sqrt{B^2-4AC}}{2A},\ \hat{p}_2 = \frac{-B+\sqrt{B^2-4AC}}{2A} \tag{2}$$

where $A = L(1-f(1-t))$, $B = -1+f-L$ and $C = 1-ft$ (Fig. 1). As shown by Hoffmann et al. (1990), when all three equilibria exist, $\hat{p}_0$ and $\hat{p}_2$ are stable while $\hat{p}_1$ between them is unstable and gives the invasion threshold (in appendix B we rederive this result extending it to $f > 1$). Turelli (1994, Eq. 4) indicated that $\hat{p}_2 > 1/2$, although they did not provide an explicit proof for that. Below, we show what is required for this result to hold.

Before proceeding to understand effects caused by $f > 1$, it is useful to understand why the classical result arises, i.e. that $f \leq 1$ predicts the absence of low and stable frequencies. We outline the reasons both mathematically and biologically. Mathematically, $\hat{p}_1$ and $\hat{p}_2$ are symmetric around a point $r = -B/2A$. Now, for $\hat{p}_2$ to remain below 1/2, it is necessary (but not sufficient) that r is below 1/2. Solving the inequality $r < 1/2$ for f with $0 < t, L \leq 1$ yields

$$\begin{aligned} \frac{1-f+L}{2L(1-f(1-t))} &< \frac{1}{2} \\ \Leftrightarrow f > \frac{1}{1-L(1-t)} &\geq 1 \end{aligned} \tag{3}$$

which clearly implies that $r > 1/2$ for any $f \leq 1$. However, the result also suggests that permitting $f > 1$ can change the results, since $r < 1/2$ becomes possible and thus possibly also $\hat{p}_2 < 1/2$. A numerical example shows this indeed happens with suitable parameter values (Figure 1d).

It is worthwhile to consider the robustness of the equilibria against stochastic variation in the environment, that may perturb populations away from their current equilibria. We show that this makes equilibria close to 0.5 unlikely to persist in the long term, within the classically assumed parameter regime where $f \leq 1$. Consider a situation where $f \leq 1$ and the stable equilibrium $\hat{p}_2$ is close to the low end of the range 0.5 to 1. We have a situation $\hat{p}_1 < r < \hat{p}_2$ but based on the assumption $f \leq 1$ we also know that $r > 0.5$. Since we assumed $\hat{p}_2$ is close to 0.5, it follows that all points $\hat{p}_1$, r and $\hat{p}_2$ must be close to each other (since r is located exactly midway between $\hat{p}_1$ and $\hat{p}_2$). The conclusion arises that if the stable equilibrium $\hat{p}_2$ is close to 0.5, then the unstable equilibrium $\hat{p}_1$ (representing the invasion barrier) is not far below $\hat{p}_2$. Under such circumstances, random fluctuations in infection prevalence can easily make the population fall stochastically under the invasion barrier, from where it deterministically proceeds towards extinction (unless another stochastic event brings it back up to above the invasion barrier).

Biologically, the above analysis implies that if *Wolbachia* infection has effects that go beyond CI, then these additional fitness differences between infected and uninfected hosts can change the outcome of the interaction significantly. The direction of these differences determine whether *Wolbachia* frequencies can only stabilize at high values (with direct negative fitness effect), or whether the range of prevalence values extends to include low frequencies.

To develop biological intuition for why this is the case, it is instructive to see why the invasion threshold exists in the classical situation ($f \leq 1$), i.e. why in Figure 1a

there is a switch from poor spread of *Wolbachia* at low frequencies to better success at moderate frequencies. The impact of *Wolbachia* on uninfected eggs is captured by parameter L, but the value of L becomes irrelevant as a determinant of the dynamics when the infection is rare (low p). If we consider a change in p, $\Delta p = p_{T+1} - p_T$, and take the limit of negligible frequencies of infected males and females in equation (1), we get $\lim_{p \to 0} \Delta p = pft - p = (ft - 1)p$.

For the infection to increase in frequency when it is rare, the requirement is that $\Delta p > 0$, creating a simple condition: $ft > 1$ means that the 'leakage' of offspring into an uninfected state (i.e., only a proportion t of offspring of *Wolbachia*-infected mothers inherit the infection) is more than compensated for by the fecundity advantage of *Wolbachia*-infected mothers. Assuming this is the case, the equilibrium $\hat{p}_0$ is unstable and *Wolbachia* can spread from rare. Simultaneously, fixation is not possible if we assume $t < 1$; the 'leakage' always maintains a supply of uninfected individuals.

Figure 1a) shows an example where ft is clearly below 1 and spread from rare is consequently not possible. When there are not enough males causing the 'competing' egg type to fail (low p), the combination of low fecundity ($f < 1$) and imperfect transmission ($t < 1$) together prevent CI-inducing *Wolbachia* from maintaining its frequency in the population. At higher p, we can no longer ignore the consequences of CI, and therefore the effect of the parameter L. Now, a high proportion of infected males will elevate the relative success of infected eggs compared to uninfected competitors; the success difference arises from uninfected females' eggs dying. This explains why in classical examples the CI-inducing symbiont can begin to spread at some threshold value of p: the spread is caused by sufficiently many infected males harming the reproduction of uninfected individuals.

The central role of the product ft provides important insight. Its value exceeds 1 when $f > 1/t$, thus sufficiently high values of f can make the invasion barrier disappear

(shown e.g. by Engelstädter and Telschow, 2009). However, this can never happen for any $f \leq 1$, since $1/t \geq 1$ for any $t \in]0, 1]$. Thus $f > 1$ is a necessary (but not sufficient) requirement for the invasion barrier to disappear. Biologically, situations of $f > 1/t > 1$ are possible, once one considers the full range of possibilities for pleiotropy. If females infected by CI-inducing *Wolbachia* are more fecund than uninfected females, and if this advantage is sufficiently large, then the invasion barrier disappears. In this case, one might at first sight expect that this additional advantage would simply make *Wolbachia* reach fixation. This is not the case, however, since the effect of $t < 1$ causing fixation to be impossible still holds.

Instead, the more interesting finding is that $f > 1$ creates conditions where a low frequency of *Wolbachia* can be stable. This was not possible when $f \leq 1$, because infected males at low prevalence do not cause a large enough survival difference between eggs. But if $f > 1$, infected females can outcompete uninfected ones with very little (assuming $ft < 1$) or no ($ft > 1$) CI-related advantages, and as a whole, the infection can spread at values of L that are so low that they would never make *Wolbachia* spread in situations involving $f \leq 1$ (Fig. 1b). In other words, the latter condition would not feature an invasion barrier above which positive frequency dependence would be sufficiently strong to bring *Wolbachia* to higher frequencies.

If we now assume that $f > 1/t$, we know that the invasion barrier is absent since infected females are more fecund than uninfected ones (as the above implies that $f > 1$). Once the infection spreads, so that infected males are no longer negligibly rare, infected females reap additional benefits (indirectly, by the effect of L killing their competitors). The remaining equilibrium in the $0 < p < 1$ range is stable. Even when $L = 0$, i.e. the effect of males on eggs is switched off and the aforementioned additional benefits are absent, qualitatively similar dynamics arise: the infection frequency increases from rare but cannot fix due to imperfect transmission (this case is fully captured by the analogy

to the asexual lineages A and B above; also shown in Fig. S1 of Zug and Hammerstein, 2018).

Figure 1b shows examples with $ft > 1$ and five values of L. As expected, the $L = 0$ case (lowest curve) exhibits simple dynamics, with the infected strain increasing until imperfect transmission leads to a stable equilibrium which, by solving for $\Delta p = pft/(pf + (1-p)) - p = 0$ is at $p = (1-ft)/(1-f) \approx 0.06$ with the given numerical example. Higher values of L lead to equilibria with a higher prevalence of *Wolbachia*. Choosing equivalent values for L in a setting where $ft < 1$ always leads to lower growth (all graphs are lower in 1c compared with the equivalent graph in 1b). At high L this means introducing an invasion barrier, while at low L the result is negative growth regardless of the current frequency of p, i.e. extinction of *Wolbachia* (two lowest curves in figure 1c).

It is worth clarifying the statement that introducing benefits can lower the equilibrium frequency of *Wolbachia*, as it is counter-intuitive at first sight. The examples show this to be true in a qualified sense: if f is the only parameter that changes, we do not expect an existing equilibrium to shift to a lower value but to a higher value. However, a high value of f can change a situation that does not permit any *Wolbachia* to persist into one where it does (the light-coloured curves in 1c; Fig. 2c). The interesting findings of equilibria with $p < 0.5$ arise because $ft > 1$ implies no invasion barrier irrespective of the value of L, thus *Wolbachia* can spread from rare even if L is meagre, but the modest value of L then gives very little 'boost' at higher values of p (meagre male effects on uninfected eggs when L is low).

Although our analysis so far appears to suggest that cases with invasion barriers associate with high equilibrium frequencies, it should be noted that examining the full range of possibilities permitted by $f > 1$ includes cases where an invasion barrier exists, and the stable equilibrium is below 0.5. An invasion threshold exists if $ft < 1$, while the

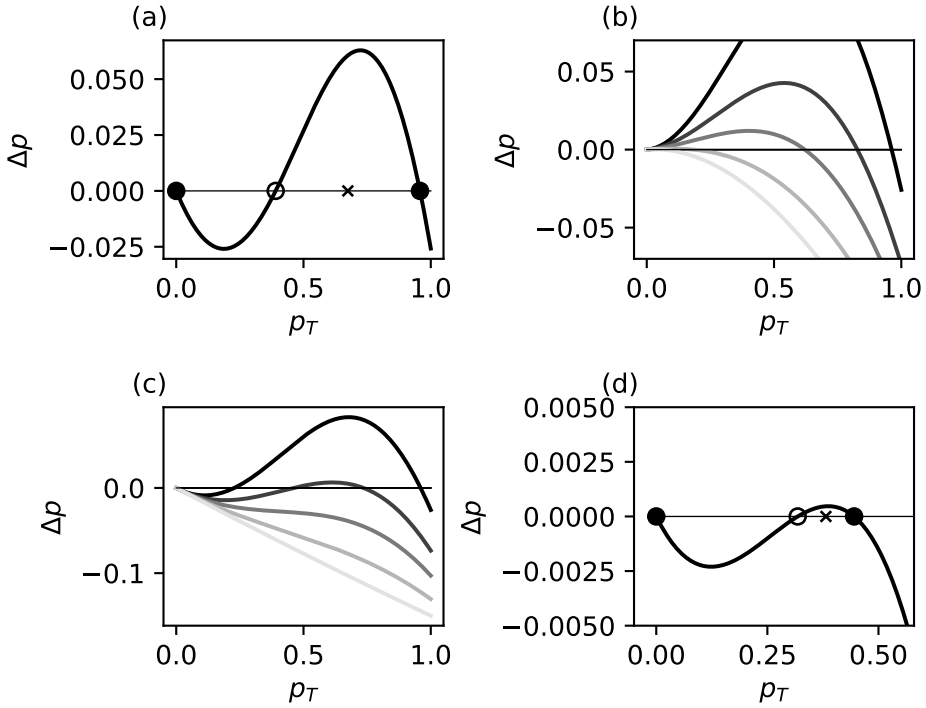

Figure 1: Equilibria of the diplodiploid model (Eq. 1). a) $\Delta p = p_{T+1} - p_T$ as a function of p_T. Dots: stable equilibria, circle: unstable equilibrium, cross: point r. Parameters $f = 0.85$, $t = 0.85$, $L = 0.85$. Examples b) and c) contrast the effect of parameter L (strength of CI) on the invasion dynamics when b) $ft > 1$ or c) $ft < 1$. Each line shows one value of L: 0.85 (darkest), 0.55, 0.35, 0.15 and 0 (lightest). In b) $f = 1.19$, $t = 0.85$, and no invasion threshold exists. Low values of L lead to low stable infection frequencies ($p < 0.5$). In c) $f = 0.99$, $t = 0.85$ ($ft < 1$). High L shows the classic case of an invasion threshold and a stable equilibrium, while low L predicts extinction of *Wolbachia*. d), it is also possible to have a low stable frequency ($p < 0.5$) together with an invasion threshold ($ft = 0.9588 < 1$) with parameters $f = 1.128$, $t = 0.85$, $L = 0.35$.

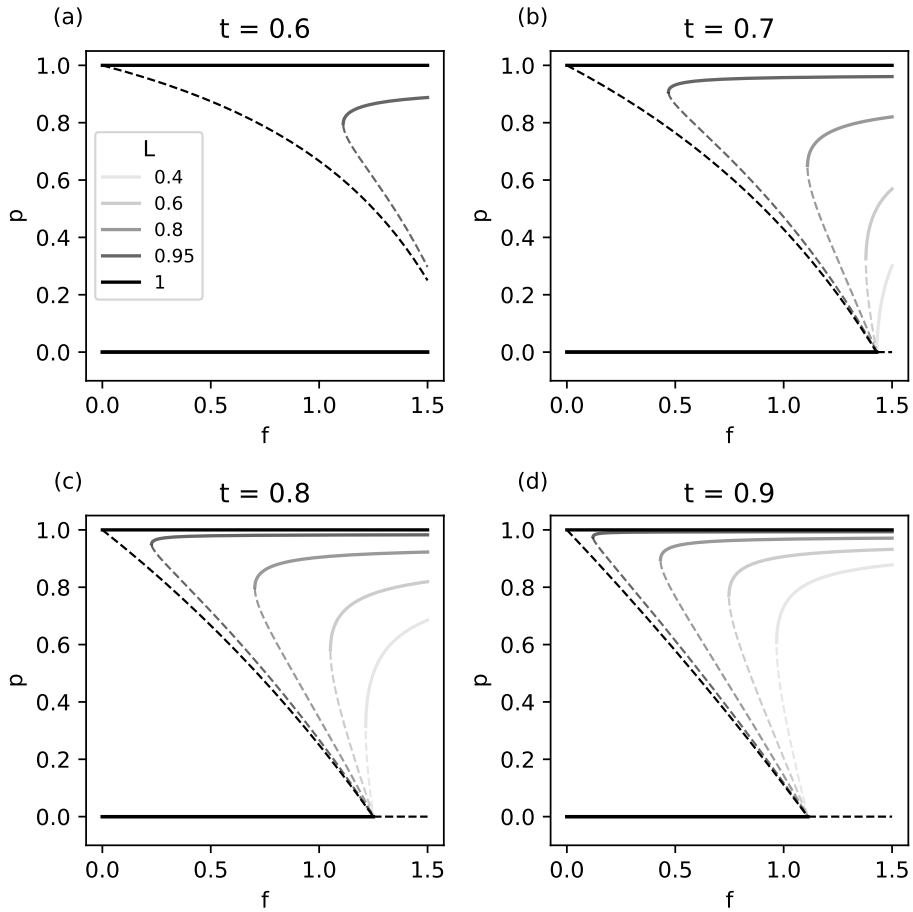

Figure 2: Equilibria of the diplodiploid model (Eq. 1) as function of f, for different values of t (panel titles) and L (colours of the lines, see legend). Solid lines: stable equilibria. Dashed lines: unstable equilibria.

necessary condition for $\hat{p}_2 < 0.5$ is that $f > 1$. Suitably chosen values for f and t can combine both criteria simultaneously (example: Figure 1d). As a summary, increasing f will increase the value of the stable equilibrium, but at the same time open possibilities for less successful symbionts, manifested as lower values of L and t, that will equilibrate at lower frequencies. Effect of parameter f with different combinations of t and L are shown in Fig. 2.

It is of interest to see if equilibria with low p can only exist at very low values of L (mild incompatibility) as in the above examples, or whether other values of other

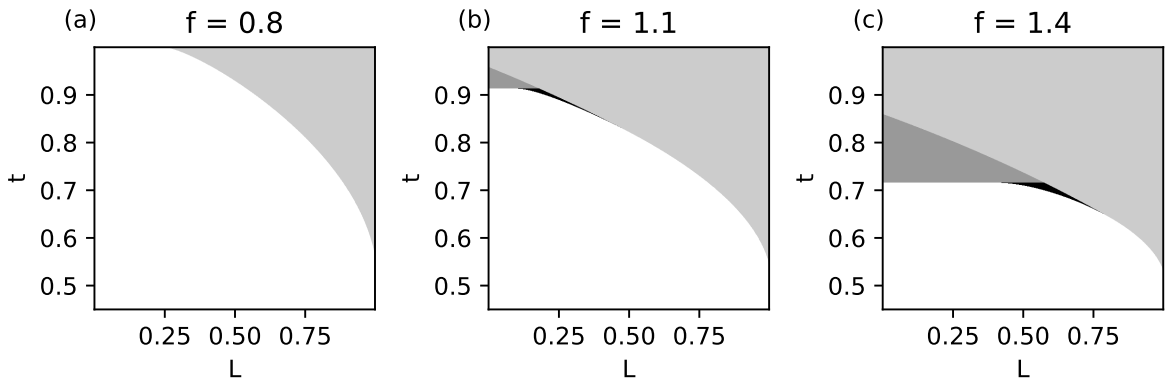

Figure 3: Equilibria of the diplodiploid model across parameter space, categorized as featuring only the trivial equilibrium at $p = 0$ (white), non-trivial stable equilibrium $\hat{p}_2 > 0.5$ (light grey), and stable equilibrium $\hat{p}_2 < 0.5$ without (dark grey) or with (black) an invasion threshold. Values of f as indicated above each panel.

parameters allow low-p equilibria to exist even if L is relatively high. The latter proves to be the case, and as a whole, assuming $f > 1$ reveals a richer set of possible outcomes than what can occur if $f \leq 1$ (Fig. 3). The "classical" example with two stable equilibria and an invasion threshold between them (light grey in Fig. 3), as well as extinction of *Wolbachia* (white), can be found whether f indicates costs or benefits to being infected, but $f > 1$ additionally permits cases without invasion barriers and with low stable prevalence of *Wolbachia* when low or moderate L combines with $ft > 1$ (dark grey areas). Finally, the situation with the stable equilibrium below 0.5 co-occurring with an invasion threshold places conflicting demands on the value of f (see above), and while not impossible to fulfil, satisfying them requires rather fine tuned parameter settings with low or intermediate values of L and for a narrow range of t slightly below $1/f$ (black areas in Fig. 3).

Haplodiploid models

Thus far, the presented analysis of low stable frequencies of *Wolbachia* assumed diplodiploid sex determination. In this section, we extend these results to two haplodiploid cases,

that differ in the effects of an infection on the host. The first case resembles the classic diplodiploid model in that cytoplasmic incompatibility kills fertilized eggs, but since under haplodiploidy all fertilized eggs develop as females, we now switch to calling it the female-killing effect (and note that it is equivalent to the "*Leptopilina* type" of Vavre et al. (2000)). In the second case, incompatibility instead leads to the loss of paternal chromosomes and haploidization of the zygote, which consequently develops as a male (Breeuwer and Werren, 1990). We call this the masculinization effect (Vavre et al. (2000) call it the "*Nasonia* type", see also Breeuwer and Werren (1990)). Note that both the female-killing effect and the masculinization effect can lead to a male-biased sex ratio, but the underlying mechanisms differ, and we model them separately.

The existence of a low-value stable equilibrium of *Wolbachia* with $f > 1$ also proves true in these two haplodiploid systems. We consider the female-killing effect first. Applying our notation from above to the model of Vavre et al. (2000), the dynamics for *Wolbachia* frequency in females and males, $p_{F,T}$ and $p_{M,T}$ respectively, become

$$\begin{aligned} p_{F,T+1} &= \frac{p_{F,T} f t}{p_{F,T} f(1-(1-t)p_{M,T}L) + (1-p_{F,T})(1-p_{M,T}L)} \\ p_{M,T+1} &= \frac{p_{F,T} f t}{p_{F,T} f + (1-p_{F,T})} \end{aligned} \tag{4}$$

For the masculinization effect, we again adopt the model from Vavre et al. (2000); the dynamics for females and males, respectively, are given as

$$\begin{aligned} p_{F,T+1} &= \frac{p_{F,T} f t}{p_{F,T} f(1-(1-t)p_{M,T}L) + (1-p_{F,T})(1-p_{M,T}L)} \\ p_{M,T+1} &= \frac{p_{F,T} f t}{p_{F,T} f(1+(1-t)\frac{k}{1-k}p_{M,T}L) + (1-p_{F,T})(1+\frac{k}{1-k}p_{M,T}L)} \end{aligned} \tag{5}$$

The female equations are identical across the two effects (Eqs. (4) and (5)), as females are lost due to infections impacting fertilized eggs in both cases. However, the male equations differ, because the masculinization effect converts females into males instead

of killing them. The magnitude of this additonal "male production" depends on the fertilization rate k. If k is high, CI has the potential to cause strong effects on the primary sex ratio, both because the baseline male production is by definition low under high k, and because there are many eggs available to be masculinized. For the female equation, the parameter k cancels out.

Appendix A gives explicit solutions for the equilibrium frequencies in these two models. There, we show analytically that the stable equilibrium frequency for *Wolbachia* infection exceeds 0.5 when $f \leq 1$, with the caveat that this applies for both sexes in the female-killing model, but only for the *Wolbachia* frequency in female hosts in the masculinization model. Regardless of this caveat, numerical examples show that the ability of $f > 1$ to create stable equilibria below 0.5 for both sexes extends to both models (examples of equilibria: Fig. 4). The proofs in Appendix A follow the same logic as for the diplodiploid model: the unstable and stable equilibria are symmetric around a point r (for each sex), which is never below 0.5 if $f \leq 1$. Hence, again, a stable equilibrium close to 0.5 may be vulnerable to the population falling under the invasion barrier if stochasticity causes fluctuations in *Wolbachia* frequency, as discussed above with the diplodiploid model.

The infection frequency in males generally does not have to be the same as in females, because of the two different ways in which males can be produced: via masculinization or via the egg not being fertilized in the first place. This additional flexibility means that the masculinization case brings about the possibility that the frequency of infection in males can stabilize below 0.5 even with $f = 0.95 < 1$ (Fig. 4c), though the parameter range for this type of situation is narrow whenever f is low (Figure 5). In all cases that we explored, the infection frequency was higher in females than in males (Fig. 6). Moreover, the female infection frequency was very high whenever $f \leq 1$, so that the average infection frequency was above 0.5 even in the few cases where male infection

frequency was below 0.5 under the masculinization model. Such cases were found with high fertilization rate (k) yielding a strongly female-biased sex ratio, hence the average infection frequency was close to that found in females. Overall, low infection frequencies seem very rare unless $f > 1$.

Software

Matlab (v. R2019b) and Mathematica (v. 12.1.1.0) were used for solving and analysing equations. The images were created and numerical analysis conducted in Python (v. 3.7.4) using packages numpy (v. 1.21.5 Harris et al., 2020) and matplotlib (v. 3.4.2 Hunter, 2007). Source code is provided in the Zenodo repository (DOI: 10.5281/zenodo.6443871).

Discussion

The traditional model of the dynamics of symbionts inducing cytoplasmic incompatibility predicts stable infection frequencies above 0.5 in diplodiploid systems (Turelli, 1994). However, the prediction of high stable frequencies is tightly linked to an assumption of direct fitness effects of infection being either neutral or detrimental. Our findings, that low-frequency stable equilibria can exist, are as such not new: examples were presented by Zug and Hammerstein (2018), but these authors did not elaborate on when exactly low frequecies can be expected to prevail. We show analytically that they can occur only if $f > 1$.

The association between negative fitness effects and high frequencies is explained by positive frequency-dependence. CI penalizes uninfected individuals' reproduction, but cannot do so very efficiently if infections are currently rare; therefore, infections must overcome an invasion barrier first but can thereafter spread to high frequencies (only counteracted by leaky transmission). Positive fitness effects remove the difficulties of

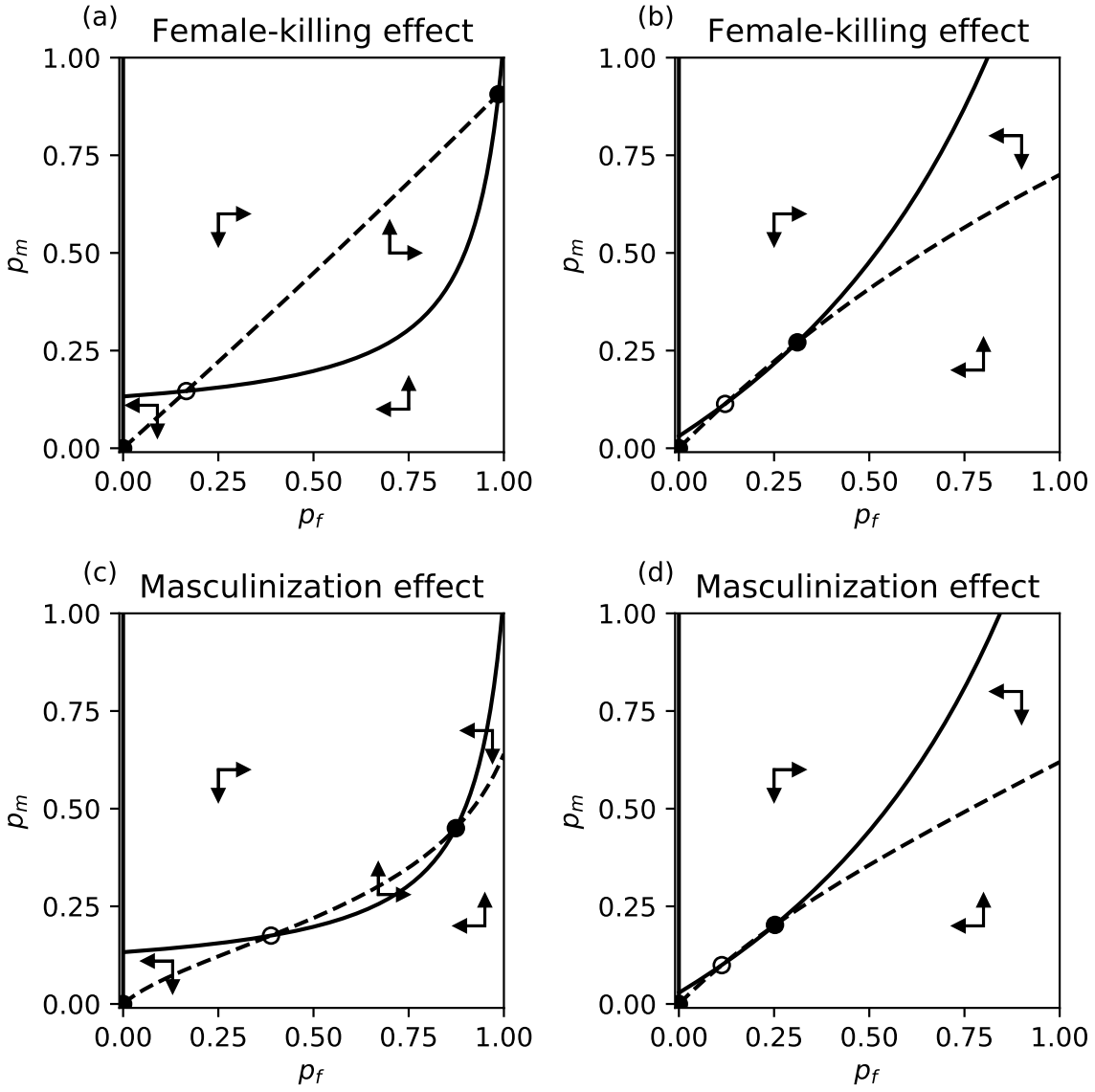

Figure 4: Equilibria of the two haplodiploid models in (p_f, p_m)-space (FK: eq. (4), MD: eq. (5)). Dots: stable equilibria, circle: unstable equilibrium. Dashed line: male null-cline ($\Delta p_m = 0$), solid line: female null-cline ($\Delta p_f = 0$). a) The stable equilibrium is close to fixation and invasion threshold exists with $f = 0.95$, $t = 0.92$, $L = 0.95$. b) Low frequency stable equilibrium with invasion threshold ($f = 1.4$, $t = 0.7$, $L = 0.65$). c) The non-trivial stable equilibrium at $\hat{p_m} < 0.5$, although $f < 1$ ($f = 0.95$, $t = 0.92$, $L = 0.95$, $k = 0.9$). d) Low frequency stable equilibrium with invasion threshold ($f = 1.4$, $t = 0.7$, $L = 0.7$, $k = 0.5$).

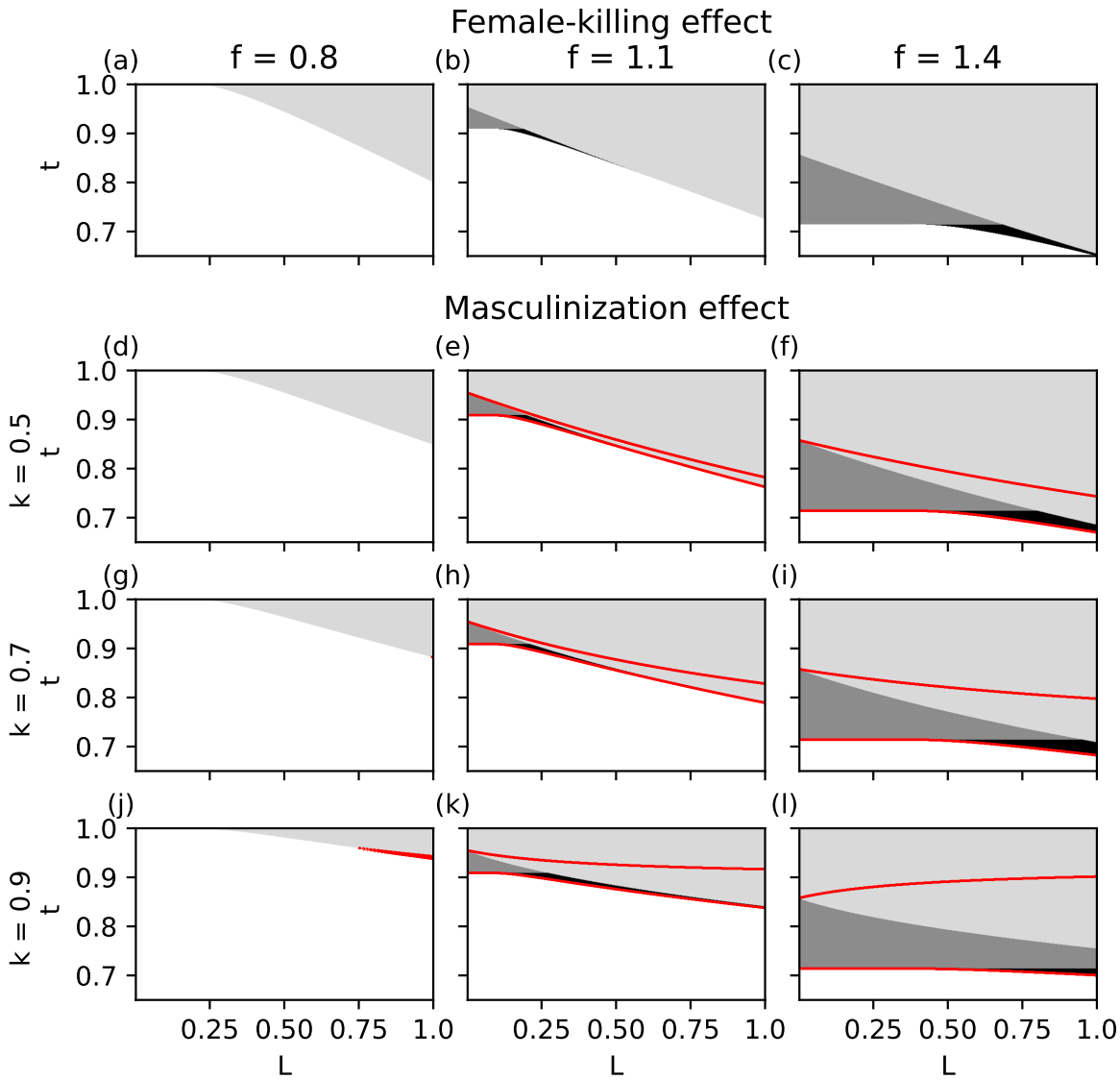

Figure 5: Classification of equilibria of the haplodiploid models across parameter space. Columns from left to right with $f = 0.8$, $f = 1.1$ and $f = 1.4$. First row: female-killing effect; other rows: masculinization effect model with different values of k. White: only the trivial equilibrium at zero. Light grey: non-trivial stable equilibrium with $0.5 < \hat{p_f} \leq 1$. Dark grey: $\hat{p_f} \leq 0.5$ without invasion threshold. Black: $\hat{p_f} \leq 0.5$ with invasion threshold. The red lines outline two thresholds, between which $\hat{p_m} \leq 0.5$. Note that the vertical t-axes span only high values, where non-trivial equilibria exist.

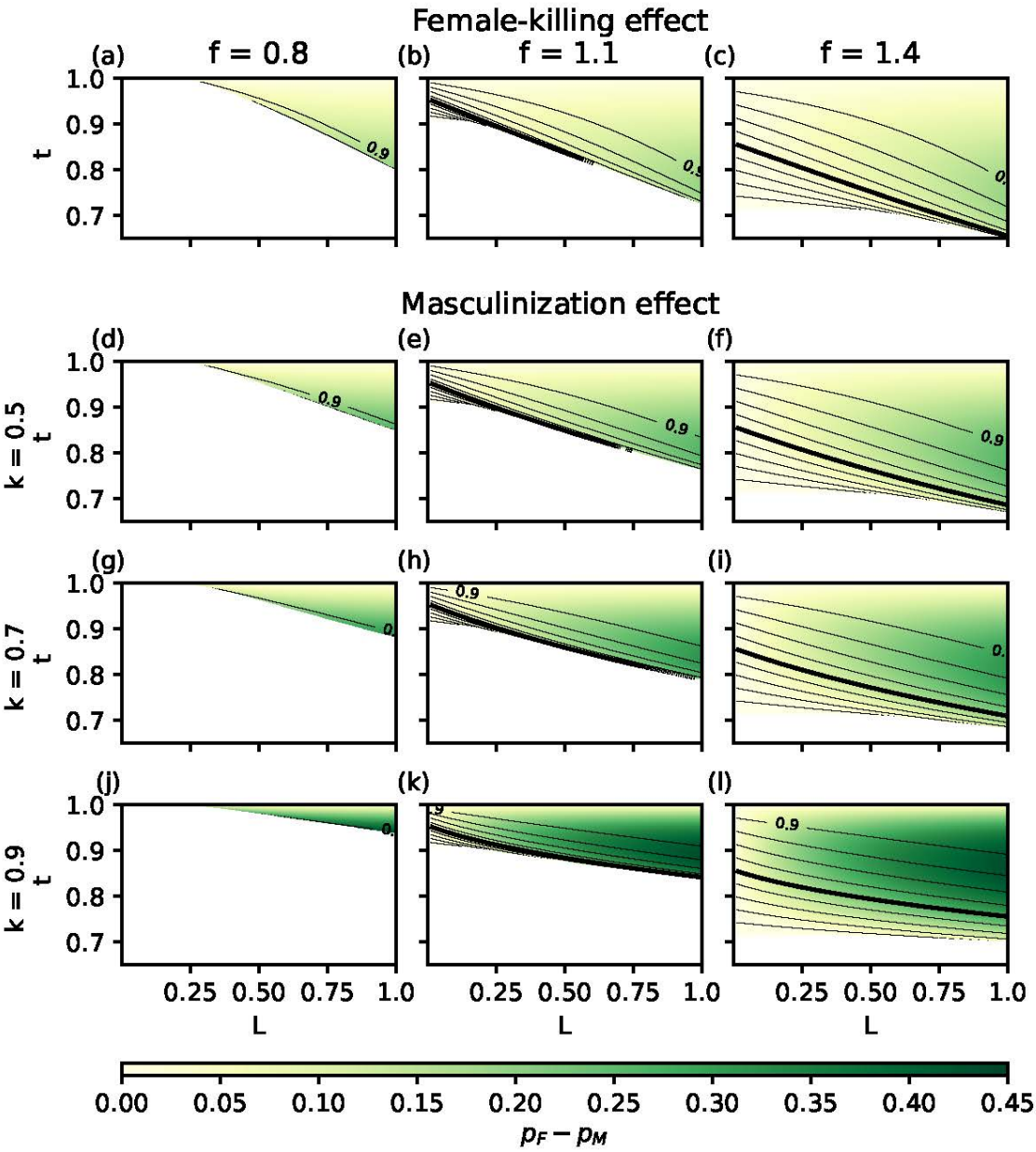

Figure 6: Equilibrium infection frequencies of the haplodiploid models across parameter space. Columns from left to right with $f = 0.8$, $f = 1.1$ and $f = 1.4$. First row: female-killing effect; other rows: masculinization effect model with different values of k. Black contour lines show infection frequency in females (with gaps of 0.1), while the color shows difference between female and male infection frequency according to the colorbar in the bottom. Thick solid contour marks 0.5 frequency. Females always have higher infection frequency than males.

spreading from very low frequencies, and as a whole this creates conditions for stable equilibria at low frequency. Similar results hold for haplodiploid systems, with the exception of male infection frequency in masculinization-type CI, which can stay below 0.5 even with $f \leq 1$.

The simplest case of low stable frequencies occurs when the effective fitness of rare infected individuals is higher than that of uninfected individuals ($ft > 1$). In that case, a rare infection can spread as there is no invasion barrier (Engelstädter and Telschow, 2009), and the infection frequency rises to levels dictated by the strength of CI, efficiency of transmission and relative fecundity. With very weak CI and leaky transmission, the stable equilibrium can be very close to zero (e.g. 0.06 in Fig. 1b). A more complex situation occurs when the relative fecundity of infected individuals exceeds one but not by much, so that $f > 1$ combines with $ft < 1$. In that case, the invasion threshold exists, but it is still possible to have a stable equilibrium below 0.5, a finding that has not been demonstrated before. As the previous analysis usually assumed $f \leq 1$, they also observed stable equilibria above 0.5.

Note that similar results arise when considering a sexual population containing a lineage with a wild-type symbiotic strain and a lineage with a fitness-increasing symbiotic strain with non-perfect transmission. As shown by Zug and Hammerstein (2018), the condition $ft > 1$ is necessary and sufficient for the successful invasion of the fitness-increasing strain, as it provides the strain with higher effective fecundity than the wild-type lineage. They also mentioned this being analogous to the mutation-selection balance of haploids (Hoffmann and Turelli, 1997). Along the same lines, when studying possibility of sex-ratio determination of CI-inducing symbionts, Egas et al. (2002) conclude that the symbiont can invade "if the proportion infected daughters produced by infected females is bigger than the proportion daughters produced by uninfected females (when mated with uninfected males)", again returning to higher effective fecundity of

the infected individuals.

Given that the consequences of $f > 1$ are not only clear but also, in the form of examples, present in earlier analyses, why have positive fitness effects not been discussed explicitly before? A typical discussion of *Wolbachia* focuses on the fitness deficit caused by infection (Hoffmann et al., 1990), which creates an interesting paradox: why should a symbiont with negative fitness effects on its host be so common in natural populations? The discovery of reproductive manipulation (such as CI) solved this conundrum. The possibility of positive fitness effects has thereafter not attracted much attention, despite the relevant equations remaining valid when $f > 1$. As an exception, Zug and Hammerstein (2018) recently explored positive fitness effects, with a focus on the evolution of the infectious agent and comparison of different symbionts (non-manipulating, CI, male-killing). There is some empirical evidence for *Wolbachia* strains that elevate host fitness, where the benefit may manifest itself, for example, as higher fecundity (Fry et al., 2004) or protection from viral diseases (Chrostek et al., 2013).

Besides showing that the stable equilibrium almost always exceeds 0.5 when $f \leq 1$, we also argue that infection frequencies slightly above 0.5 are not expected to be robust in nature. Our model analysis thus suggests that positive fitness effects of CI-inducing *Wolbachia* infection are expected to be found in those natural populations that show low and stable densities of infection over time. Stable infection frequencies below 0.5 and even slightly above it are likely paired with positive fitness effect unless another dynamic can explain the situation. Besides the fitness benefit demonstrated here, low frequencies may of course be observable during ongoing invasion or result from source-sink dynamics.

It would be ideal to complement our findings with concrete empirical examples of stable CI-inducing *Wolbachia* strains that are maintained at low frequencies in their host populations. Unfortunately, it is difficult to highlight one specific case that perfectly illustrates our findings. Although it is well appreciated that *Wolbachia* is common among

species while not necessarily reaching high frequencies within a species (Sazama et al., 2019), studies typically do not provide much information regarding additional fitness effects when populations have been screened for *Wolbachia* infection (Russell et al., 2012). Often, there is no information on whether the symbiont induces CI or any other phenotype in their hosts (Duplouy et al., 2015). The stability properties of reported low frequencies are rarely evaluated, as most studies report prevalence of the symbiont at a single time point; recurrent screenings of the same natural populations over the course of several generations are rare (however, see Kriesner et al., 2013; Duplouy et al., 2015, 2021). Hence, the changes in the infection frequencies through generations remain unknown, and it is difficult to determine stability of the infection frequency, let alone any existence of invasion barriers or other dynamical features of the system.

Our model highlights that when strains are found at low frequencies in their host populations, there are additional possibilities to them being either in the process of being eliminated or of invading (Duplouy et al., 2010; Kriesner et al., 2013) their host population: a low frequency can be a stable outcome. The question about the state of potentially ongoing change can partially be answered with the study of associated mtDNA genotype diversity, but often even this data is not available (but see Duplouy et al., 2010; Richardson et al., 2012). Long-term surveys of host populations and their infection status through time, as well as the study of their ecology are therefore needed. The recent clear results from the effective control strategy against human-born disease such as Dengue have clearly shown that long-term surveys are possible and informative for endosymbionts (eg. Ryan et al., 2019).

Acknowledgements

We are thankful for helpful comments of Roman Zug and Martin Kapun on the earlier versions of this work. Funding was provided by the Swiss National Foundation to HK,

PK and CdV. CdV was also supported by an Academy of Finland grant (no. 340130, awarded to Jussi Lehtonen). AD was funded by Academy of Finland grant number 321543.

Appendix A: Limits for CI frequency in haplodiploid systems

Female-killing effect

When CI causes mortality of the female embryo in haplodiploid species, the population dynamics follow the model in Eq. (4). The non-trivial equilibria of the model are

$$\hat{p}_F = \frac{2 - 2f - ft + f^2t + fLt \pm ft\sqrt{f^2 + (L-1)^2 + f(-2 - 2L + 4Lt)}}{2\left[1 + f(-2 + f + Lt + fLt(t-1))\right]}$$

$$\hat{p}_M = \frac{1 - f + L \pm \sqrt{f^2 + (L-1)^2 + f(-2 - 2L + 4Lt)}}{2L} \tag{A1}$$

where pluses belong to the non-trivial stable equilibrium $\hat{p}_2$ and minuses to the unstable equilibrium $\hat{p}_1$ when they exist (Vavre et al., 2000). The stability of these is explored numerically in Appendix B.

We use the same technique as for diplodiploid model, showing that the real part r (excluding the square root part) of the equilibrium frequency is below 0.5 if and only if $f > 1$, and hence $\hat{p}_2 > 0.5$ for $f \leq 1$. For the male frequency, it is simple to see that

$$r_M = \frac{1 - f + L}{2L} < \frac{1}{2} \Leftrightarrow 1 - f < 0 \Leftrightarrow f > 1$$

but for the female frequency the proof is a bit more complicated. In that case

$$r_F = \frac{2 - 2f - ft + f^2t + fLt}{2\left[1 + f(-2 + f + Lt + fLt(t-1))\right]}$$

The denominator of r_F divided by two (i.e. the square brackets) is a quadratic function

of f:

$$(1 + Lt(t-1))f^2 + (Lt-2)f + 1,$$

which is always positive, as the intercept is positive and no real roots exist (discriminator $Lt^2(L-4) < 0$). As the denominator of r_F is always larger than zero, we can multiply the inequation $r_F < 1/2$ with the denominator, resulting in an equivalent inequation

$$(1-t)(1-Lt)f^2 + tf - 1 > 0 \tag{A2}$$

In eq. (A2) the intercept of the left hand side is negative and the quadratic coefficient $(1-t)(1-Lt) > 0$. Hence, the equation holds true for positive f only when f is larger than the larger root, say f_{lim}, of the quadratic function (A2). Only if $f_{lim} < 1$, can $f \leq 1$ fulfil the inequality (A2) and equivalently $r_F < 1/2$. Let us now investigate in what conditions $f_{lim} < 1$.

$$\begin{aligned}
f_{lim} = \frac{-t + \sqrt{t^2 + 4(1-t)(1-Lt)}}{2(1-t)(1-Lt)} &< 1 \\
\Leftrightarrow \sqrt{t^2 + 4(1-t)(1-Lt)} &< 2(1-t)(1-Lt) + t \quad |\,()^2\text{, both sides positive} \\
\Leftrightarrow t^2 + 4(1-t)(1-Lt) &< 4(1-t)^2(1-Lt)^2 + 4t(1-t)(1-Lt) + t^2 \\
\Leftrightarrow 4(1-t)^2(1-Lt) &< 4(1-t)^2(1-Lt)^2 \quad |\,: 4(1-t)^2(1-Lt) > 0 \\
\Leftrightarrow 1 &< 1 - Lt \\
\Leftrightarrow Lt &< 0
\end{aligned}$$

Thus, the lower limit f_{lim} is below one only if $Lt < 0$, which is never true. Hence the limit is always at least 1, which implies that $r_F < 1/2$ is true only if $f > 1$. This concludes the proof that $\hat{p}_{F,2} > 1/2$ for the FK-type CI when $f \leq 1$.

Masculinization effect

The CI can alternatively lead to deletion of the paternal genome, turning the diploid egg into haploid that develops as male. In that case, the population dynamics follow the model in equation (5). The non-trivial equilibria of the model are

$$\begin{aligned}\hat{p}_F &= \frac{2+f^2(1+k)t-f(2+(1+k+(-1+k)L)t)}{2(1-f(2+(-1+k)Lt+f(-1+(-1+k)L(-1+t)t))} \\ &\pm \frac{ft\sqrt{(-1+k)((-1+k)(1-f+L)^2+4L(-1+k+fk(-1+t))(-1+ft))}}{2(1-f(2+(-1+k)Lt+f(-1+(-1+k)L(-1+t)t))} \\ \hat{p}_M &= \frac{1-f-k+fk+L-kL}{2L(1-k+fk(1-t))} \\ &\pm \frac{\sqrt{(-1+k)((-1+k)(1-f+L)^2+4L(-1+k+fk(-1+t))(-1+ft))}}{2L(1-k+fk(1-t))}\end{aligned} \tag{A3}$$

where pluses belong to the non-trivial stable equilibrium $\hat{p}_2$ and minuses to the unstable equilibrium $\hat{p}_1$ when they exist (Vavre et al., 2000). The stability of these is explored numerically in Appendix B.

The real part of the solution for female frequency is

$$r_F = \frac{2+f^2(1+k)t-f(2+(1+k+(-1+k)L)t)}{2(1-f(2+(-1+k)Lt+f(-1+(-1+k)L(-1+t)t))}.$$

The denominator of r_F is a linear function of k

$$2-4f+2f^2+2fLt+2f^2L(-1+t)t+k(-2fLt-2f^2L(-1+t)t)$$

The slope of that function is $-2fLt(1+f(t-1))<0$ and the root is

$$1+\frac{(f-1)^2}{fLt(1+f(t-1))} \geq 1,$$

which means, that the denominator is positive for $k < 1$, and the inequality $r_F < 1/2$ simplifies to equal condition

$$1 - f^2 - ft + f^2t + f^2Lt - f^2Lt^2 + k(-ft + f^2t + f^2L(-1+t)t) < 0. \tag{A4}$$

The left hand side is again a linear function of k, which has an intercept

$$\begin{aligned} &1 - f^2 - ft + f^2t + f^2Lt - f^2Lt^2 \\ &= (1-f)(1+f-ft) + f^2t(1-t)L \end{aligned}$$

that is always positive for $L \in]0,1]$ as it is linear in L and has positive intercept and slope. The linear function of k in Eq. (A4) has one root at

$$\hat{k} = 1 + \frac{1 - f^2 - 2f(1-f)t}{ft(1 - f + fL(1-t))} > 1$$

as the second term is always positive. The denominator of the second term is obviously positive and the numerator is also positive, as it is linear in t, having negative slope and root $(1+f)/2f \geq 1$ when $f \leq 1$.

Thus, (A4) has positive intercept and the root $\hat{k} \geq 1$, which means that the left hand side of (A4) is positive for all $k \in]0,1[$ and hence the inequality (A4) is never true for valid values of k. As A4 is equivalent with $r_F < 1/2$, this implies that the real part of the female equilibrium frequency is never less than 1/2 with valid parameter values (assuming $f \leq 1$). The actual stable equilibrium is even higher than r_F, which concludes the proof that female frequency is never below 1/2 at the stable equilibrium.

Finally, for the infection frequency in males under the masculinization effect model, it is easy to prove by example that the frequency at the stable equilibrium indeed can be below 1/2 even with $f < 1$. In Figure 4c, the parameters are $f = 0.95$, $t = 0.92$, $L = 0.95$, $k = 0.9$ while $\hat{p}_{M,2} < 1/2$.

In conclusion, the infection frequency in females follows the same pattern as FK-type and diplodiploid systems with symmetry of the stable equilibrium and the unstable equilibrium around a point that is above 0.5 for all $f \leq 1$, but the infection frequency in males makes an exception as it can go below 0.5 even with $f < 1$

Appendix B: Stability analysis of the equilibria of CI models

In this appendix we will use linear stability analysis to calculate the stability of the equilibria of the different models presented in the main text.

Diplodiploid model

The frequency of *Wolbachia* in both males and females in this model is given by equation 1 in the main text, here repeated for convenience,

$$p_{T+1} = \frac{p_T f t}{p_T f(1-(1-t)p_T L) + (1-p_T)(1-p_T L)}. \tag{B1}$$

For notational convenience, define a function g as

$$g(p_T) = \frac{f t}{p_T f(1-(1-t)p_T L) + (1-p_T)(1-p_T L)},$$

so that

$$p_{T+1} = g(p_T)p_T. \tag{B2}$$

Equilibria of this discrete time dynamical system satisfy

$$\hat{p} = g(\hat{p})\hat{p}. \tag{B3}$$

An equilibrium, $\hat{p}$, is locally stable if and only if the absolute value of the derivative of p_{T+1} with respect to p_T is smaller than one, that is,

$$\left| \left. \frac{\mathrm{d}p_{T+1}}{\mathrm{d}p_T} \right|_{\hat{p}} \right| < 1.$$

We will denote this derivative evaluated at equilibrium $\hat{p}$ as $\zeta(\hat{p})$.

Therefore to calculate the stability of the equilibria of our diplodiploid model, we need the derivative of p_{T+1} with respect to p_T. Differentiate equation (B2) to obtain

$$\frac{\mathrm{d}p_{T+1}}{\mathrm{d}p_T} = g(p_T) + p_T \frac{\mathrm{d}g}{\mathrm{d}p_T}.$$

Next we want to evaluate this derivative at the equilibria. Equation B1 has three equilibria, solved from equation $Ap^2 + Bp + C = 0$, as written in equation (2) in the main text, here repeated for convenience:

$$\hat{p}_0 = 0,\ \hat{p}_1 = \frac{-B - \sqrt{B^2 - 4AC}}{2A},\ \hat{p}_2 = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \tag{B4}$$

where $A = L(1 - f(1 - t))$, $B = -1 + f - L$ and $C = 1 - ft$.

The stability of the first equilibrium, $\hat{p}_0$, is given by

$$\begin{aligned} \zeta(\hat{p}_0) &= \left| \left. \frac{\mathrm{d}p_{T+1}}{\mathrm{d}p_T} \right|_{\hat{p}_0} \right|, \\ &= g(0), \\ &= ft. \end{aligned}$$

The equilibrium $\hat{p}_0 = 0$ is therefore stable if and only if $ft < 1$. At $ft = 1$, the equilibria $\hat{p}_0$ and $\hat{p}_1$ coincide, and when $ft > 1$ $\hat{p}_0$ is unstable.

The stability of the second equilibrium, $\hat{p}_1$, is given by

$$\zeta(\hat{p}_1) = \left| g(\hat{p}_1) + \hat{p}_1 \left. \frac{\mathrm{d}g}{\mathrm{d}p_T} \right|_{\hat{p}_1} \right|,$$

The first term $g(\hat{p}_1) = 1$ when $\hat{p}_1 \neq 0$, because of equation (B3). The second term is

$$\begin{aligned}
\hat{p}_1 \frac{\mathrm{d}g}{\mathrm{d}p_T}\bigg|_{\hat{p}_1} &= \frac{-\hat{p}_1}{ft}\left(f - 1 - L + 2\hat{p}_1 L\left(1 - f(1-t)\right)\right) \\
&= -\frac{\hat{p}_1}{ft}\left(B + 2\hat{p}_1 A\right),
\end{aligned}$$

therefore

$$\zeta(\hat{p}_1) = \left|1 - \frac{\hat{p}_1}{ft}\left(B + 2\hat{p}_1 A\right)\right|,$$

or using equation (B4),

$$\zeta(\hat{p}_1) = \left|1 + \frac{\hat{p}_1}{ft}\sqrt{B^2 - 4AC}\right|.$$

Stability requires that $\zeta(\hat{p}_1) < 1$. Since $ft > 0$ always, equilibrium $\hat{p}_1$ is unstable whenever $\hat{p}_1 > 0$. The first non-trivial equilibrium is therefore always unstable when it is positive, in agreement with Hoffmann et al. (1990).

Similarly, the stability of the second equilibrium, $\hat{p}_2$, is given by

$$\begin{aligned}
\zeta(\hat{p}_2) &= \left|g(\hat{p}_2) + \hat{p}_2 \frac{\mathrm{d}g}{\mathrm{d}p_T}\bigg|_{\hat{p}_2}\right| \\
&= \left|1 - \frac{\hat{p}_2}{ft}\sqrt{B^2 - 4AC}\right|.
\end{aligned}$$

The equilibrium $\hat{p}_2$ is therefore stable if and only if

$$0 < \frac{\hat{p}_2}{ft}\sqrt{B^2 - 4AC} < 2. \tag{B5}$$

If $\hat{p}_2 > 0$, then the left-hand side of this inequality is true. To check the right-hand side

of the inequality, we substitute the expression of $\hat{p}_2$ in equation (B4) into (B5) to obtain:

$$\frac{B^2 - 4AC - B\sqrt{B^2 - 4AC}}{2A(1-C)} < 2,$$

or equivalently

$$B\left(\frac{B - \sqrt{B^2 - 4AC}}{4A}\right) < 1,$$

which can be written as

$$\frac{-B}{2}\hat{p}_2 < 1. \tag{B6}$$

Now, if $B < 0$ (i.e. $f < 1 + L$) we can multiply with $-2/B$ to get an equivalent stability condition

$$\hat{p}_2 < \frac{2}{1 + L - f},$$

which is always true since $0 < 1 + L - f < 2$ and the right-hand side is larger than 1. Therefore as long as $\hat{p}_2 \leq 1$ this condition is fulfilled and $\hat{p}_2$ is a stable equilibrium. On the other hand, if $B \geq 0$ ($f \geq 1 + L$), the left hand side of ineq. (B6) is zero or negative since $-B < 0$ and the inequality is true for all $\hat{p}_2 \geq 0$. Hence, $\hat{p}_2$ is always stable if it exists within the biologically relevant range between 0 and 1.

Haplodiploid models

The *Wolbachia* models in the main text are all frequency-dependent models that in general we can write as follows

$$\mathbf{p}_{T+1} = \mathbf{B}(\mathbf{p}_T)\mathbf{p}_T, \tag{B7}$$

where $\mathbf{p}_T$ is a vector with the frequency of *Wolbachia* infection in females and males as its entries. For the diplodiploid model, frequencies in females and males are identical, and we can therefore model the frequency in the entire population using only a scalar,

p_T.

Equilibria of this discrete time dynamical system satisfy

$$\hat{\mathbf{p}} = \mathbf{B}(\hat{\mathbf{p}})\hat{\mathbf{p}}.$$

An equilibrium, $\hat{\mathbf{p}}$, is locally stable if and only if the absolute value of the derivative of $\mathbf{p}_{T+1}$ with respect to $\mathbf{p}_T$ is smaller than one, that is,

$$\left| \left. \frac{\mathrm{d}\mathbf{p}_{T+1}}{\mathrm{d}\mathbf{p}_T^{\intercal}} \right|_{\hat{\mathbf{p}}} \right| < 1.$$

The left-hand side of this equation is called the Jacobian matrix. In general, the stability of an equilibrium to small perturbations can be determined by the dominant eigenvalue of the Jacobian matrix of the linearization of equation (B7) at the boundary equilibrium. We will denote the dominant eigenvalues by ζ.

The Jacobian matrix

$$\mathbf{M} = \left. \frac{\mathrm{d}\mathbf{p}_{T+1}}{\mathrm{d}\mathbf{p}_T^{\intercal}} \right|_{\hat{\mathbf{p}}}$$

is obtained by differentiating equation (B7) and evaluating the resulting derivative at the equilibrium of interest. We will carry out this differentiation using matrix calculus; for details of the methodology see Magnus and Neudecker (1985, 1988); for ecological presentations see for example Caswell (2007). Matrix calculus makes extensive use of the vec $\cdot$ operator, which stacks the columns of a matrix; e.g.,

$$\mathrm{vec} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$$

and of the fact that

$$\operatorname{vec} \mathbf{ABC} = (\mathbf{C}^{\intercal} \otimes \mathbf{A}) \operatorname{vec} \mathbf{B}$$

for any conformable matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, where $\otimes$ is the Kronecker product.

Differentiating equation (B7) gives

$$\mathrm{d}\mathbf{p}_{T+1} = \mathbf{B}\mathrm{d}\mathbf{p}_T + (\mathrm{d}\mathbf{B})\, \mathbf{p}_T,$$

where the explicit dependence of $\mathbf{B}$ on $\mathbf{p}$ has been omitted to avoid a cluttering of brackets. Apply the vec operator to both sides and remember that the matrix $\mathbf{B}$ in our case is of size 2×2, to obtain

$$\mathrm{d}\mathbf{p}_{T+1} = \mathbf{B}\mathrm{d}\mathbf{p}_T + (\mathbf{p}_T^{\intercal} \otimes \mathbf{I}_2)\, \mathrm{dvec}\mathbf{B}.$$

Thus

$$\begin{aligned} \mathbf{M} &= \left.\frac{\mathrm{d}\mathbf{p}_{T+1}}{\mathrm{d}\mathbf{p}_T^{\intercal}}\right|_{\hat{\mathbf{p}}} \\ &= \mathbf{B}(\hat{\mathbf{p}}) + (\hat{\mathbf{p}}^{\intercal} \otimes \mathbf{I}_2) \left.\frac{\partial \mathrm{vec}\mathbf{B}(\mathbf{p})}{\partial \mathbf{p}^{\intercal}}\right|_{\hat{\mathbf{p}}}. \end{aligned}$$

Both the female killing model, and the masculinization effect model have the following general shape:

$$\begin{pmatrix} p_F \\ p_M \end{pmatrix}_{T+1} = \begin{pmatrix} g(p_{F,T}, p_{M,T}) & 0 \\ h(p_{F,T}, p_{M,T}) & 0 \end{pmatrix} \begin{pmatrix} p_F \\ p_M \end{pmatrix}_T,$$

such that

$$\mathbf{B}(\mathbf{p}) = \begin{pmatrix} g(p_{F,T}, p_{M,T}) & 0 \\ h(p_{F,T}, p_{M,T}) & 0 \end{pmatrix},$$

and

$$\begin{aligned}\mathbf{M} &= \mathbf{B}(\hat{\mathbf{p}}) + (\hat{\mathbf{p}}^{\intercal} \otimes \mathbf{I}_2)\frac{\partial \mathrm{vec}\mathbf{B}(\mathbf{p})}{\partial \mathbf{p}^{\intercal}}\bigg|_{\hat{\mathbf{p}}} \\ &= \begin{pmatrix} g(\hat{\mathbf{p}}) & 0 \\ h(\hat{\mathbf{p}}) & 0 \end{pmatrix} + \hat{p}_F \begin{pmatrix} \frac{\partial g}{\partial p_F} & \frac{\partial g}{\partial p_M} \\ \frac{\partial h}{\partial p_F} & \frac{\partial h}{\partial p_M} \end{pmatrix}\bigg|_{\hat{\mathbf{p}}}\end{aligned} \tag{B8}$$

Female-killing effect

We consider the female-killing effect first, the model repeated here for convenience (equation 4 from the main text):

$$\begin{aligned} p_{F,Tt+1} &= \frac{p_{F,T} f t}{p_{F,T} f(1-(1-t)p_{M,T}L) + (1-p_{F,T})(1-p_{M,T}L)}, \\ p_{M,T+1} &= \frac{p_{F,T} f t}{p_{F,T} f + (1-p_{F,T})}. \end{aligned}$$

For the trivial equilibrium, $\mathbf{p}_0 = (0,0)$, $g(\hat{\mathbf{p}}) = h(\hat{\mathbf{p}}) = ft$

$$\mathbf{M} = \begin{pmatrix} ft & 0 \\ ft & 0 \end{pmatrix}.$$

The eigenvalues of this Jacobian matrix are 0 with eigenvector $(p_F, p_M) = (0,1)$, and ft with eigenvector $(p_F, p_M) = (1,1)$. The zero eigenvalue is associated with perturbations in the directions of males only, which makes biological sense (these perturbations decay back to the zero equilibrium since a population of males only will go extinct). The largest eigenvalue of the $\mathbf{p}_0 = (0,0)$ equilibrium is ft. Similarly to the haplodiploid model, the trivial equilibrium is therefore stable if $ft < 1$, and unstable if $ft > 1$.

The non-trivial equilibria of this model are given by the solution in equation (A1),

$$\hat{p}_F = \frac{2 - 2f - ft + f^2t + fLt \pm ft\sqrt{f^2 + (L-1)^2 + f(-2 - 2L + 4Lt)}}{2\left[1 + f(-2 + f + Lt + fLt(t-1))\right]}$$

$$\hat{p}_M = \frac{1 - f + L \pm \sqrt{f^2 + (L-1)^2 + f(-2 - 2L + 4Lt)}}{2L}$$

To calculate the Jacobian matrix for the non-trivial equilibria, we need to calculate the derivatives in equation (B8). We can rewrite the term under the square roots in equilibria above as follows:

$$f^2 + (L-1)^2 + f(-2 - 2L + 4Lt) = (1 + L - f)^2 + 4L(ft - 1),$$

and therefore

$$\hat{p}_M = \frac{1 + L - f \pm \sqrt{(1 + L - f)^2 + 4L(ft - 1)}}{2L}.$$

This rewrite makes it clear that when $ft > 1$, $\hat{p}_M$ of the smaller equilibrium is negative, and when $ft = 1$, $\hat{p}_M = 0$ for the smaller equilibrium.

It will be useful to note that for the non-trivial equilibria, $\hat{\mathbf{p}} = \mathbf{B}(\hat{\mathbf{p}})\hat{\mathbf{p}}$ implies that $g(\hat{p}_F, \hat{p}_M) = 1$, and $h(\hat{p}_F) = \frac{\hat{p}_M}{\hat{p}_F}$. It follows that

$$\begin{aligned} \frac{\partial h}{\partial p_F} &= \frac{1-f}{ft}\frac{\hat{p}_M^2}{\hat{p}_F^2}, \\ \frac{\partial h}{\partial p_M} &= 0, \end{aligned}$$

and

$$\frac{\partial g}{\partial p_F} = -\frac{1}{ft} g(p_F, p_M)^2 \left[f\left(1 - (1-t)\, p_M L - 1 + p_M L\right)\right]$$

and therefore

$$\left.\frac{\partial g}{\partial p_F}\right|_{\hat{\mathbf{p}}} = \frac{1}{ft}\left[1-\hat{p}_M L - f\left(1-(1-t)\,\hat{p}_M L\right)\right]. \tag{B9}$$

The equilibrium condition $g(\hat{p}_F, \hat{p}_M) = 1$ can be rewritten as

$$\hat{p}_F f\left(1-(1-t)\,\hat{p}_M L\right) = ft - 1 + \hat{p}_M L + \hat{p}_F - \hat{p}_F \hat{p}_M L.$$

Multiplying equation (B9) with $\hat{p}_F$ as in equation (B8), and substituting the above expression yields

$$\left.\frac{\partial g}{\partial p_F}\right|_{\hat{\mathbf{p}}} \hat{p}_F = \frac{1}{ft}\left[1-\hat{p}_M L - ft\right].$$

Similarly,

$$\frac{\partial g}{\partial p_M} = \frac{1}{ft} g(p_F, p_M)^2 \left[p_F f(1-t)L - L(1-p_F)\right],$$

and

$$\begin{aligned}
\left.\frac{\partial g}{\partial p_M}\right|_{\hat{\mathbf{p}}} &= \frac{1}{ft}\left[\hat{p}_F f(1-t)L + L(1-\hat{p}_F)\right], \\
&= \frac{1}{\hat{p}_M ft}\left[\hat{p}_F f + (1-\hat{p}_F) - ft\right], \\
&= \frac{1}{\hat{p}_M}\left[\frac{\hat{p}_F}{\hat{p}_M} - 1\right],
\end{aligned}$$

where I have used $g(\hat{p}_F, \hat{p}_M) = 1$ and $h(\hat{p}_F, \hat{p}_M) = \frac{\hat{p}_M}{\hat{p}_F}$ for the last two lines, respectively.

Putting all of this together, we get the following Jacobian matrix,

$$\begin{aligned}
\mathbf{M} &= \begin{pmatrix} 1 & 0 \\ \frac{\hat{p}_M}{\hat{p}_F} & 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{ft}\left(1-\hat{p}_M L - ft\right) & \frac{\hat{p}_F}{\hat{p}_M}\left[\frac{\hat{p}_F}{\hat{p}_M}-1\right] \\ \frac{(1-f)}{ft}\frac{\hat{p}_M^2}{\hat{p}_F} & 0 \end{pmatrix}, \\
&= \begin{pmatrix} \frac{1}{ft}\left(1-\hat{p}_M L\right) & \frac{\hat{p}_F}{\hat{p}_M}\left[\frac{\hat{p}_F}{\hat{p}_M}-1\right] \\ \frac{\hat{p}_M}{\hat{p}_F}\left[1+\frac{(1-f)}{ft}\hat{p}_M\right] & 0 \end{pmatrix},
\end{aligned}$$

the eigenvalues of this Jacobian are given by the following quadratic equation

$$\lambda^2 - \frac{\lambda}{ft}\left(1-\hat{p}_M L\right) - \left(1+\frac{1-f}{ft}\hat{p}_M\right)\left(\frac{\hat{p}_F}{\hat{p}_M}-1\right) = 0$$

If the absolute value of the dominant eigenvalue of the Jacobian is below one, then the equilibrium is stable. The eigenvalues are given by

$$\lambda = \frac{1}{2ft}\left(1-\hat{p}_M L\right) \pm \frac{1}{2ft}\sqrt{\left(1-\hat{p}_M L\right)^2 + 4(ft)^2 C},$$

where

$$C = \left(1+\frac{1-f}{ft}\hat{p}_M\right)\left(\frac{\hat{p}_F}{\hat{p}_M}-1\right).$$

Since $L \leq 1$, the term before square root is positive and the eigenvalue with the largest absolute value is

$$\lambda_1 = \frac{1}{2ft}\left(1-\hat{p}_M L\right) + \frac{1}{2ft}\sqrt{\left(1-\hat{p}_M L\right)^2 + 4(ft)^2 C}.$$

We could not determine stability of the equilibria analytically. Instead, we used Python to explore numerically the parameter space $0.01 \leq L, t \leq 1 \times 0.01 \leq f \leq 3$ sampling uniformly 100 values for each parameter (limits included), which resulted in 1 000 000 parameter combinations. For all cases where $\lambda_1 \neq 1$ (i.e. linear stability analysis did not fail) with $0 < \hat{p}_M, \hat{p}_F \leq 1$ the linear stability analysis concluded that $\hat{p}_1$ was

unstable and $\hat{p}_2$ was stable.

Masculinization effect

For the masculinization effect, the dynamics for females and males, respectively, are given by equation (5) in the main text,

$$p_{F,T+1} = \frac{p_{F,T} f t}{p_{F,T} f(1-(1-t)p_{M,T}L) + (1-p_{F,T})(1-p_{M,T}L)}$$
$$p_{M,T+1} = \frac{p_{F,T} f t}{p_{F,T} f(1+(1-t)\frac{k}{1-k}p_{M,T}L) + (1-p_{F,T})(1+\frac{k}{1-k}p_{M,T}L)}$$

The non-trivial equilibria of this model are given by equation (A3):

$$\hat{p}_F = \frac{2+f^2(1+k)t - f(2+(1+k+(-1+k)L)t)}{2(1-f(2+(-1+k)Lt+f(-1+(-1+k)L(-1+t)t))} \pm \frac{ft\sqrt{(-1+k)((-1+k)(1-f+L)^2+4L(-1+k+fk(-1+t))(-1+ft))}}{2(1-f(2+(-1+k)Lt+f(-1+(-1+k)L(-1+t)t))}$$

$$\hat{p}_M = \frac{1-f-k+fk+L-kL}{2L(1-k+fk(1-t))} \pm \frac{\sqrt{(-1+k)((-1+k)(1-f+L)^2+4L(-1+k+fk(-1+t))(-1+ft))}}{2L(1-k+fk(1-t))}.$$

As before, at the trivial equilibrium, $\mathbf{p}_0 = (0,0)$, $g(\hat{\mathbf{p}}) = h(\hat{\mathbf{p}}) = ft$, and therefore

$$\mathbf{M} = \begin{pmatrix} ft & 0 \\ ft & 0 \end{pmatrix}.$$

Again the eigenvalues of this Jacobian matrix are 0 with eigenvector $(p_F, p_M) = (0,1)$, and ft with eigenvector $(p_F, p_M) = (1,1)$. The zero eigenvalue is associated with perturbations in the directions of males only. The largest eigenvalue of the $\mathbf{p}_0 = (0,0)$ equilibrium is ft. Similarly to the haplodiploid model, the trivial equilibrium is there-

fore stable if $ft < 1$, and unstable if $ft > 1$.

For the non-trivial equilibrium, we need to calculate the derivatives in the formula for the Jacobian again (equation (B8)):

$$\begin{aligned}
\left.\frac{\partial g}{\partial p_F}\right|_{\hat{\mathbf{p}}} &= \frac{1}{ft}\left[1 - \hat{p}_M L + f(1-t)\hat{p}_M L - f\right], \\
\left.\frac{\partial g}{\partial p_M}\right|_{\hat{\mathbf{p}}} &= \frac{1}{ft} L\left[1 - \hat{p}_F + \hat{p}_F f(1-t)\right], \\
\left.\frac{\partial h}{\partial p_F}\right|_{\hat{\mathbf{p}}} &= \frac{1}{ft}\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \left[1 + \frac{k}{1-k}\hat{p}_M L - f - (1-t)\frac{k}{1-k} f \hat{p}_M L\right], \\
\left.\frac{\partial h}{\partial p_M}\right|_{\hat{\mathbf{p}}} &= \frac{1}{ft}\frac{kL}{1-k}\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \left[\hat{p}_F - \hat{p}_F f(1-t) - 1\right].
\end{aligned}$$

Using $g(\hat{p}_F, \hat{p}_M) = 1$ and $h(\hat{p}_F, \hat{p}_M) = \frac{\hat{p}_M}{\hat{p}_F}$, we can simplify

$$\begin{aligned}
\left.\frac{\partial g}{\partial p_F}\right|_{\hat{\mathbf{p}}} \hat{p}_F &= \frac{1}{ft}\left[1 - ft - \hat{p}_M L\right], \\
\left.\frac{\partial h}{\partial p_F}\right|_{\hat{\mathbf{p}}} \hat{p}_F &= \frac{1}{ft}\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \left[1 - \frac{\hat{p}_F}{\hat{p}_M} ft + \frac{k}{1-k}\hat{p}_M L\right].
\end{aligned}$$

Putting it all together gives the following expression for the Jacobian matrix,

$$\mathbf{M} = \begin{pmatrix} \frac{1}{ft}\left(1 - \hat{p}_M L\right) & \frac{1}{ft} L\hat{p}_F\left(1 - \hat{p}_F + \hat{p}_F f(1-t)\right) \\ \frac{1}{ft}\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \left(1 + \frac{k}{1-k}\hat{p}_M L\right) & \frac{1}{ft}\frac{k}{1-k}\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 L\hat{p}_F\left(\hat{p}_F - \hat{p}_F f(1-t) - 1\right) \end{pmatrix}.$$

The eigenvalues of this Jacobian are given by the solution to

$$\lambda^2 - \lambda\frac{1}{ft}\left[1 - \hat{p}_M L - \left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \frac{kB}{1-k}\right] - \frac{1}{(ft)^2}\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \frac{B}{1-k} = 0,$$

where

$$B = L\hat{p}_F\left(1 - \hat{p}_F + \hat{p}_F f(1-t)\right).$$

The eigenvalues are therefore equal to

$$\lambda = \frac{1}{2ft}\left[1 - \hat{p}_M L - \left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \frac{kB}{1-k}\right] \pm \frac{1}{2ft}\sqrt{\left[1 - \hat{p}_M L - \left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \frac{kB}{1-k}\right]^2 + 4\left(\frac{\hat{p}_M}{\hat{p}_F}\right)^2 \frac{B}{1-k}}$$

As with the female killing model, we could not determine the stability of the equilibria analytically. Instead, we conducted similar numerical analysis in Python with 100 values for each $0.01 \leq L, t \leq 1$, $0.01 \leq f \leq 3$ and with three values of $k = 0.05, 0.7, 0.9$, resulting in 3 000 000 parameter combinations in total. We calculated the values of the equilibria and determined their stability based on the larger absolute value of the eigenvalues above, whenever $0 < \hat{p}_M, \hat{p}_F \leq 1$. Again, the equilibria were unstable ($\hat{p}_1$) and stable ($\hat{p}_2$) as expected, whenever they occurred in biologically meaningful range and the leading eigenvalue had absolute value different from 1.

Literature Cited

Arthofer, W., M. Riegler, D. Schneider, M. Krammer, W. J. Miller, and C. Stauffer. 2009. Hidden Wolbachia diversity in field populations of the European cherry fruit fly, Rhagoletis cerasi (Diptera, Tephritidae). Molecular Ecology **18**:3816–3830.

Breeuwer, J. A., and J. H. Werren. 1990. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. Nature **346**:558–560.

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

Caswell, H. 2007. Sensitivity analysis of transient population dynamics. Ecology letters **10**:1–15.

Chrostek, E., M. S. Marialva, S. S. Esteves, L. A. Weinert, J. Martinez, F. M. Jiggins, and L. Teixeira. 2013. Wolbachia variants induce differential protection to viruses in Drosophila melanogaster: a phenotypic and phylogenomic analysis. PLoS Genet **9**:e1003896.

Deng, J., G. Assandri, P. Chauhan, R. Futahashi, A. Galimberti, B. Hansson, L. Lancaster, Y. Takahashi, E. I. Svensson, and A. Duplouy. 2021. Wolbachia-driven selective sweep in a range expanding insect species,. Research Square .

Duplouy, A., and O. Brattström. 2018. Wolbachia in the Genus Bicyclus: a Forgotten Player. Microbial ecology **75**:255–263.

Duplouy, A., C. Couchoux, I. Hanski, and S. van Nouhuys. 2015. Wolbachia infection in a natural parasitoid wasp population. PloS one **10**:e0134843.

Duplouy, A., G. D. Hurst, S. L. O'NEILL, and S. Charlat. 2010. Rapid spread of male-killing Wolbachia in the butterfly Hypolimnas bolina. Journal of evolutionary biology **23**:231–235.

Duplouy, A., A. Nair, T. Nyman, and S. van Nouhuys. 2021. Long-term spatio-temporal genetic structure of an accidental parasitoid introduction, and local changes in prevalence of its associated Wolbachia symbiont. Authorea Preprints .

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

Engelstädter, J., and A. Telschow. 2009. Cytoplasmic incompatibility and host population structure. Heredity **103**:196–207.

Fine, P. E. 1978. On the dynamics of symbiote-dependent cytoplasmic incompatibility in culicine mosquitoes. Journal of Invertebrate Pathology **31**:10–18.

Flor, M., P. Hammerstein, and A. Telschow. 2007. Wolbachia-induced unidirectional cytoplasmic incompatibility and the stability of infection polymorphism in parapatric host populations. Journal of Evolutionary Biology **20**:696–706.

Fry, A. J., M. R. Palmer, and D. M. Rand. 2004. Variable fitness effects of Wolbachia infection in Drosophila melanogaster. Heredity **93**:379–389.

Harris, C. R., K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. 2020. Array programming with NumPy. Nature **585**:357–362.

Hilgenboecker, K., P. Hammerstein, P. Schlattmann, A. Telschow, and J. H. Werren. 2008. How many species are infected with Wolbachia?–a statistical analysis of current data. FEMS microbiology letters **281**:215–220.

Hoffmann, A. A., and M. Turelli, 1997. Cytoplasmic incompatibility ininsects. Pages 42–80 *in* S. L. O'Neill, A. A. Hoffmann, and J. H. Werren, editors. Influential passengers. Inherited microorganisms and arthropodreproduction. Oxford University Press.

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

Hughes, G., P. Allsopp, S. Brumbley, M. Woolfit, E. McGraw, and S. L. O'Neill. 2011. Variable infection frequency and high diversity of multiple strains of Wolbachia pipientis in Perkinsiella planthoppers. Applied and environmental microbiology **77**:2165–2168.

Hunter, J. D. 2007. Matplotlib: A 2D graphics environment. Computing in science & engineering **9**:90–95.

Jeong, G., T. Kang, H. Park, J. Choi, S. Hwang, W. Kim, Y. Choi, K. Lee, I. Park, H. Sim, et al. 2009. Wolbachia infection in the Korean endemic firefly, Luciola unmunsana (Coleoptera: Lampyridae). Journal of Asia-Pacific Entomology **12**:33–36.

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 Pathog **9**:e1003607.

Laven, H. 1956. Cytoplasmic inheritance in Culex. Nature **177**:141–142.

Lehtonen, J., and H. Kokko. 2012. Positive feedback and alternative stable states in inbreeding, cooperation, sex roles and other evolutionary processes. Philosophical Transactions of the Royal Society B: Biological Sciences **367**:211–221.

Magnus, J. R., and H. Neudecker. 1985. Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. Journal of Mathematical Psychology **29**:474–492.

Magnus, J. R., and H. Neudecker. 1988. Matrix differential calculus with applications in statistics and econometrics. Wiley.

Richardson, M. F., L. A. Weinert, J. J. Welch, R. S. Linheiro, M. M. Magwire, F. M. Jiggins, and C. M. Bergman. 2012. Population genomics of the Wolbachia endosymbiont in Drosophila melanogaster. PLoS genetics **8**:e1003129.

Russell, J. A., C. F. Funaro, Y. M. Giraldo, B. Goldman-Huertas, D. Suh, D. J. Kronauer, C. S. Moreau, and N. E. Pierce. 2012. A veritable menagerie of heritable

bacteria from ants, butterflies, and beyond: broad molecular surveys and a systematic review. PLoS One **7**:e51027.

Ryan, P. A., A. P. Turley, G. Wilson, T. P. Hurst, K. Retzki, J. Brown-Kenyon, L. Hodgson, N. Kenny, H. Cook, B. L. Montgomery, et al. 2019. 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 research **3**.

Sazama, E. J., S. P. Ouellette, and J. S. Wesner. 2019. Bacterial endosymbionts are common among, but not necessarily within, insect species. Environmental entomology **48**:127–133.

Sun, X., L. Cui, and Z. Li. 2007. Diversity and phylogeny of Wolbachia infecting Bactrocera dorsalis (Diptera: Tephritidae) populations from China. Environmental entomology **36**:1283–1289.

Tagami, Y., and K. Miura. 2004. Distribution and prevalence of Wolbachia in Japanese populations of Lepidoptera. Insect molecular biology **13**:359–364.

Telschow, A., M. Flor, Y. Kobayashi, P. Hammerstein, and J. H. Werren. 2007. Wolbachia-induced unidirectional cytoplasmic incompatibility and speciation: mainland-island model. PLoS one **2**:e701.

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

Vavre, F., F. Fleury, J. Varaldi, P. Fouillet, and M. Bouleatreau. 2000. Evidence for female mortality in Wolbachia-mediated cytoplasmic incompatibility in haplodiploid insects: epidemiologic and evolutionary consequences. Evolution **54**:191–200.

Weinert, L. A., E. V. Araujo-Jnr, M. Z. Ahmed, and J. J. Welch. 2015. The incidence of bacterial endosymbionts in terrestrial arthropods. Proceedings of the Royal Society B: Biological Sciences **282**:20150249.

Yen, J. H., and A. R. Barr. 1971. New hypothesis of the cause of cytoplasmic incompatibility in Culex pipiens L. Nature **232**:657–658.

Zug, R., and P. Hammerstein. 2012. Still a host of hosts for Wolbachia: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PloS one **7**:e38544.

Zug, R., and P. Hammerstein. 2018. Evolution of reproductive parasites with direct fitness benefits. Heredity **120**:266–281.