Predicting distributions of Wolbachia strains through host ecological contact – who's manipulating whom?

Clive Darwell[1], Daniel Souto-Vilaros[2], Jan Michalek[3], Sotiria Boutsi[4], Brus Iusa[5], Mentap Sisol[6], George Weiblen[7], Vojtech Novotny[8], and Simon Segar[4]

[1]National Science and Technology Development Agency
[2]Charles University Faculty of Science
[3]Czech Academy of Sciences
[4]Harper Adams University College Department of Crop and Environment Sciences
[5]New Guinea Binatang Research Center
[6]Binatang Research Center
[7]University of Minnesota
[8]Institute of Entomology, Czech Academy of Sciences

March 23, 2021

Abstract

Barriers to gene-flow within populations, typically in response to divergent selection, are often mediated via third-party interactions. Under these conditions speciation is inextricably linked to ecological context. We present a novel framework for studying arthropod speciation as mediated by *Wolbachia*, a microbial endosymbiont capable of causing host cytoplasmic incompatibility (CI) via alternative strain associations. Building on empirical findings, our model predicts that sympatric host sister-species harbour paraphyletic strains that provide CI, while well-defined congeners in ecological contact and recently diverged noninteracting congeners are uninfected due to *Wolbachia* redundancy. We argue that *Wolbachia* may provide an adaptive advantage when coupled with reduced hybrid fitness (via trait mismatching), by facilitating assortative mating between co-occurring divergent phenotypes – the *contact contingency* hypothesis. To test this, we applied a custom-built predictive algorithm to empirical data from host-specific pollinating fig wasps, achieving [?]88.46% accuracy. We then considered post-zygotic offspring mortality during CI matings by developing a model featuring fitness clines across oviposition resources. This *oviposition trade-off* model, tested through simulation, favoured CI at realistic conspecific mating frequencies despite fecundity losses. We demonstrate that a rules-based algorithm accurately predicts *Wolbachia* infection status. This has implications among other systems where closely-related sympatric species encounter adaptive disadvantage through hybridisation.

Introduction

Recognising the conditions that favour speciation is critical if we are to understand the extent and structure of biodiversity. Moreover, species interactions, both between and within trophic levels, can be significant contributors to diversification processes and are sculpted by evolutionary forces, which in combination with abiotic drivers deliver an ecosystem or community's (dynamic) state (Harmon *et al.* 2019). Thus, an in depth understanding of adaptive processes alongside their ecological contingencies (e.g., interaction strengths and polarities; Segar *et al.* 2020) is a fundamental component of the diagnostic tool kit essential for achieving standard objectives in ecology.

Apropos of this, increasing emphasis on the arthropod microbiome as a modifier of ecological interaction strength (e.g., Hansen & Moran 2014) underlines the need to consider endosymbionts as part of the extended

1

phenotype. The rise and fall of microbial partners is an eco-evolutionary process, driving and being driven by ecological interactions between organisms and their environment. One such endosymbiotic bacterium, *Wolbachia* , infects up to 40% of arthropod species and often plays a key role in speciation (Werren*et al.* 2008; Zug & Hammerstein 2015). *Wolbachia* commonly induces cytoplasmic incompatibility (CI) via sexual sterility between infected males and females that are either uninfected (*unidirectional* CI) or carry an alternative strain (*bidirectional* CI) (Beckmann *et al.* 2017, 2019; LePage*et al.* 2017). CI may therefore promote reproductive isolation (RI) between populations or incipient host species and increase the speed or likelihood of speciation by restricting geneflow (Bordenstein*et al.* 2001; Zimmer 2001; Telschow *et al.* 2007), a critical factor in sympatric and ecological speciation otherwise caused by correlations between divergent traits, mate choice and/or habitat choice (Gavrilets 2004).

In some non-arthropod taxa, *Wolbachia* is an essential mutualist and accordingly shows strict co-divergence with hosts (Casiraghi*et al.* 2001; Balvín *et al.* 2018). Among arthropods, *Wolbachia* lineages are mostly facultative and evolutionarily unstable symbionts generally exhibiting host co-phylogenetic incongruence (Shoemaker *et al.* 2002; Yang *et al.* 2012; Jäckel *et al.* 2013; Zug & Hammerstein 2015), although exceptions are known where essential mutualism appears likely (e.g., Dedeine*et al.* 2001; Raychoudhury *et al.* 2009; Hamm *et al.*2014). At broader taxonomic scales (e.g., families, orders), a non-random distribution of *Wolbachia* has been noted (Engelstädter & Hurst 2006; Weinert *et al.* 2015), viewed as the consequence of accelerated host switching among closely related species from highly speciose clades (Engelstädter & Hurst 2006). However, at reduced scales *Wolbachia* often appears idiosyncratically distributed (Shoemaker*et al.* 2002; Smith *et al.* 2012; Yang *et al.* 2012; Jäckel *et al.* 2013; Zug & Hammerstein 2015), as closely related hosts often harbour paraphyletic strains. These strains are paraphyletic within the context of the complete *Wolbachia* phylogeny, such that grouping strains from two closely related hosts renders them paraphyletic. Horizontal exchange also occurs between unrelated species (Shoemaker *et al.* 2002; Zug & Hammerstein 2015; Bailly-Bechet*et al.* 2017). Counterintuitively, this may not be readily predicted from close ecological contact (Haine & Cook 2005; Jäckel*et al.* 2013; Gerth *et al.* 2013) but incidences where it has been recorded (Sintupachee *et al.* 2006; McFrederick & Rehan 2016; Miraldo & Duplouy 2019) suggest that outcomes may be context dependant. Many studies conclude that infection status depends on the ability of *Wolbachia* to manipulate its arthropod hosts (Werren*et al.* 2008; Zug & Hammerstein 2015), which may add to the sense that it is non-systemically distributed. Testable models linking eco-evolutionary processes to distribution patterns and ecological context remain critically absent.

As *Wolbachia* mediated CI results in post-zygotic mortality, initial fitness losses due to reduced fecundity are costly, meaning that selection may be expected to operate on hosts to purge *Wolbachia* . However, *Wolbachia* is posited to facilitate reproductive isolation between incipient species in combination with reduced hybrid fitness, even when only unidirectional pre-zygotic isolation operates (Shoemaker *et al.* 1999). The maladaptation of intermediate forms is central to models of sympatric/ecological speciation which may be likely under bi-directional CI as documented in closely-related, co-occurring *Nasonia* wasps (Bordenstein & Werren 2007). Thus, it is possible that *Wolbachia* represents a tolerable cost (contingent on ecological circumstances), rendering host fitness advantage (i.e., via hybrid avoidance) the prime determinant of infection status rather than the bacterium's manipulative capability.

Predictive phylogenetic models of *Wolbachia* distribution have not previously incorporated the intensity of ecological contact between insect lineages that (a), provides a direct opportunity for horizontal exchange of microbes or genetic material, and (b), provides a contingency axis of whether RI is required. When speciation occurs in allopatry, specific mechanisms of RI may not necessarily evolve as the nascent species are not in contact (Coyne & Orr 2004). This may also be true if newly formed species specialise on different resources in sympatry (Nosil 2012). However, a mechanism of pre- or post-mating RI is required if ecological contact occurs, when the species use the same resources and overlap in space and time (e.g., Via & Hawthorne 2002).

Wolbachia typically drops out of host lineages after approximately 7 million years ($\pm$5.2-9.6) (Bailly-Bechet *et al.*2017), contributing to the lack of correlated host-symbiont divergence and adding weight to the idea that purging may occur. Compared with *Wolbachia* , alternative mechanisms of RI that require cytogenetic

2

or morphological modification may take longer periods of time to evolve (Bordenstein *et al.* 2001), and thus may not be responsive enough to changing ecological circumstances that favour diversification, particularly in a sympatric setting. These lines of evidence suggest that observed lineage dropout (Bailly-Bechet *et al.* 2017) may result from temporal changes in the relative adaptive benefits of *Wolbachia* (as alternative mechanisms of RI evolve), that may subsequently become redundant and eradicated if hosts can mediate their own infection statuses (e.g., via physiological immune responses) – hereafter termed the *adaptive decay* hypothesis.

Fig wasps (Chalcidoidea), where *Wolbachia* prevalence is ca. 60%, appear to be a prime candidate for CI manipulation because many closely related and often cryptic species (both pollinating and non-pollinating) share an enclosed reproductive space (i.e., fig inflorescences), where they regularly come into contact giving potential for hybridisation (Molbo *et al.* 2003; Darwell *et al.* 2014; Yu *et al.*2019). Moreover, inbreeding is also common favouring female biased sex-ratios, strain fidelity through vertical transmission, and reduced allospecific (i.e., *ex* community) encounter rates – all increasing barriers to gene flow (Branca *et al.* 2009). Due to the confined nature of fig syconia (i.e., fig inflorescences), co-occurring incipient species must rapidly employ RI barriers (Nosil 2012) to avoid any hybridisation costs. Fig wasp studies often show paraphyletic *Wolbachia* infections across sister-species (Shoemaker *et al.* 2002; Haine & Cook 2005; Yang *et al.* 2012), while species occupying fig communities that do not contain congeners invariably display negative *Wolbachia* statuses (Haine & Cook 2005). Importantly, while these factors may obviously and measurably dominate fig wasp community structure, their influence may be apparent in all ecological systems to differing degrees.

Hybridisation between highly adapted lineages of wasps, with narrow abiotic niches and extreme matching for host fig interacting traits, presents a rather extreme cost. We develop a model which selects for ecologically contingent host tolerance of otherwise costly *Wolbachia* in this system, thus imposing pre-zygotic selection and reduced gene flow between lineages. We propose that sister populations/incipient species of wasps, associated with diverging fig hosts, should be infected with paraphyletic *Wolbachia* strains when in close ecological contact. Thus, *Wolbachia* should facilitate adaptive divergence. Subsequently, we model purging of*Wolbachia* after alternative mechanisms of RI are established across evolutionary time (see Fig. 1). This *contact contingency* hypothesis leads to a predictive system that would elicit an apparently stochastic distribution, with respect to the host phylogeny, similar to those frequently observed.

While the unusual ecological conditions of fig wasps, including the potential for complementary pre-zygotic (e.g., behavioural) barriers, may be sufficient to permit tolerance of post-zygotic fecundity reduction, we also develop a second model to singularly account for post-zygotic dynamics. We consider the heightened value of oviposition sites which are finite for pollinating fig wasps as they are unable to leave fig syconia after entry (Cook & Segar 2010). In monoecious fig species syconium oviposition sites are more valuable towards the centre where parasitoid wasp ovipositors typically do not penetrate (e.g., Dunn*et al.* 2008). As intermediate hybrid forms exhibit marked fitness reductions within co-evolved systems, the costs of reduced fecundity may prove tolerable if hybrid eggs are not wasted on premium oviposition sites. This could feasibly occur in two ways among fig wasps: (i) via preferential oviposition of favoured non-hybrid embryos (Hymenoptera are at least known to manipulate the oviposition order of haploid versus diploid eggs as well as adjust sex ratios; Raja *et al.* 2008); or (ii) via differential mortality affecting unviable hybrids before oviposition (an undocumented but plausible phenomenon). This is contingent on multiple mating events occurring within syconia (e.g., Murray 1990; Greeff *et al.* 2003), so that fig wasp foundresses carry egg loads of high versus low fitness embryos.

We model this *oviposition trade-off* hypothesis by simulating pre-oviposition egg mortality causing reduced egg load, meaning zero fitness is attributed to lost hybrid embryos. However, as fig wasps are known to prioritise oviposition into favourable sites, remaining non-hybrid eggs receive greater average fitness as they are probabilistically oviposited towards the syconium centre as opposed to their average position when mixed together with viable hybrids of reduced fitness (Fig. 2). We examine these trade-offs under different frequencies of conspecific-heterospecific mating opportunity scenarios.

As part of an ongoing study investigating patterns of co-speciation between several monophyletic fig (*Fi-*

cus , Moraceae) species complexes and their pollinating wasps, we tested our primary 'contact contingency' hypothesis, which explains how *Wolbachia* infection may be an adaptive responses to diversification pressures in host wasps. This is accomplished by presentation of empirical data of pollinating wasps screened for *Wolbachia* , and then using Python programming to simulate our proposed mechanism incorporating ecological contact and phylogenetic relationships. We then test our secondary 'oviposition trade-off' hypotheses which accounts for *Wolbachia* post-zygotic fecundity costs by modelling disparities in oviposition site quality again using Python programming. Our field site, located along a steep elevational gradient in Papua New Guinea, features a steep clinal turnover of *Ficus* species (Segar *et al.* 2017) with species complexes comprising lowland/highland sister species or morphologically homogenous species with wide elevational ranges (Souto-Vilarós *et al.* 2018, 2019).

Methods

Field collection

We collected pollinating fig wasps from one species complex (*Ficus itoana* , *F. umbrae* and *F. microdyctia*) two sub-species (*F. trichocerasa* subsp. *trichocerasa* and *F. trichocerasa* subsp. *pleioclada*) and two species with wide elevational ranges (*F. wassa* and *F. arfakensis*). For each species we collected near-ripe syconia (enclosed inflorescences) from 10-15 fig individuals at each elevation and placed these into breathable rearing pots to allow for the emergence of pollinating fig wasps. A selection of five male and five female wasps were then stored in 2ml tubes three-quarters filled with silica gel and a small piece of cotton wool before being transferred to -20°C for long term storage (Moe & Weiblen 2012).

DNA extraction and PCR

Wasp samples in open tubes were submerged in liquid nitrogen and manually homogenized with sterile plastic pestle. Subsequent DNA extractions were performed using DNeasy Blood & Tissue kits (Qiagen) following several modifications of the manufacturer's protocol. The initial lysis step was done gently at 37°C overnight and finished by 30 minutes incubation with 1 µL of RNase A (Qiagen) per sample. To enhance the yield the final elution step was done with total of 200 µL of deionized water separated in two rounds of column washing. The resulting DNA solution was then dried using vacuum concentrator and resuspended in 38 µL of buffer EB in order to increase the concentration. One microliter was used for Qubit quantification, 2 µL were loaded on 2% agarose gel pre-stained with GelRed® Nucleic Acid Gel Stain (Biotium) and ran at 120V for 70 minutes to assess the quality of extracted DNA. Samples were quantified using a Qubit 3 Fluorometer (ThermoFisher Scientific) and diluted to a total of ˜20 ng in 35 µL of EB buffer (Qiagen). We used the primers and protocols of Baldo et al. (Baldo *et al.* 2006) to amplify the *Wolbachia* surface protein gene (*wsp*) and the five Multi Locus Strain Typing (MLST) genes used for accurate strain typing and better detection of recombination. All PCR products were sequenced using Sanger sequencing, chromatograms were checked for multiple peaks and edited in Finch TV (Geospiza: https://digitalworldbiology.com/FinchTV) before alignment using BioEdit (Hall 1999).

Strain typing and molecular phylogenies

All *wsp* and MLST sequences were compared to those in the MLST data base to assess strain similarity. All sequences were assigned to the nearest matching allele. Final strain delimitation was based on i) consistency of allele assignation across MLST loci and *wsp* and ii) phylogenetic evidence. Separate phylogenetic trees were generated for each gene, including all existing MLST strains and *wsp* , to further assess the consistency of strain groupings across genes and verify that no two strains consistently formed a monophyletic group within the wider context of the MLST data base. The sequences were trimmed using BioEdit (Hall, 1999) and aligned using ClustalW (default settings) as implemented in BioEdit. All alignments were manually checked and edited and all MLST loci gained from each individual sample in this study were concatenated using Seaview (Gouy *et al.* 2010). Single gene trees were computed using RAxML (Stamatakis 2014) (CAT model) and the concatenated multi-gene phylogeny was estimated using ExaBayes (default settings for one run, majority rules consensus using a threshold of 50%). The phylogeny of pollinating wasps was estimated using genomic data taken from Souto-Vilarós et al. (2019) using ExaBayes (Aberer *et al.* 2014) as outlined

4

above. The computation of phylogenetic analyses was performed employing resources of CIPRES Science Gateway (Miller *et al.* 2015).

Simulation of Wolbachia *distribution among host species*

We used Python programming to model our 'contact contingency' hypothesis and its proposed mechanism incorporating ecological contact and *Wolbachia* purging in a phylogenetic context (wolPredictor: https://github.com/ctdarwell/wolPredictor using the *wolPredictor_xmeansDelim.py* script). Our program divides fig wasp samples into putative *species clusters* according to co-phenetic branch-length distances (Sneath & Sokal 1973) via an inputted phylogeny using X-means in the '*pyclustering*' Python library (Novikov 2019). X-means derived species clusters were derived incrementally for species richness cut-offs between two (default minimum) and 100 (greater than a credible upper limit). wolPredictor then assigns *Wolbachia* infection statuses according to rules based on our working hypothesis (Fig. 1). Briefly: (i) across communities, distinct *Wolbachia* strains are assigned to each different putative fig wasp species cluster co-occurring in the same fig-host community (thus, our model represents bi-directional CI), unless only a single putative species occurs within a community upon which a non-infection is scored for these wasps; (ii) at each species clustering iteration, *Wolbachia* is purged according to incrementally increasing (starting from no purging) cut-off thresholds between putative species clusters within a community; and (iii) arbitrarily assigned strain names [i.e., from (i) and modified in (ii)] across all purging scenarios are then conservatively matched to the empirically derived *wsp* clade strains and recorded (see github documentation for a detailed breakdown of wolPredictor functionality including scripts provided for investigation of outputted files). We ran our analyses 20 times to account for variation in X-means species delimitation outcomes. Predictive success for the range of putative species delimitation clusters was then assessed by statistical comparison against a distribution derived from 1000 random predictions across the dataset and also from a set of 100 iterations of wolPredictor with predictions generated after randomly shuffling *wsp* clade associations.

Simulation of fig wasp inclusive fitness outcomes under differential hybrid embryo mortalities

We again used Python programming to model our 'oviposition trade-off' hypothesis (https://github.com/ctdarwell/wolPredictor/ciFitnessModel.py) and investigate how offspring fitness might benefit from preferential oviposition facilitated by reduced foundress egg load (or via selective oviposition) despite losses in fecundity resulting from the post-zygotic mechanism of CI. Our model allocates each foundress 1000 eggs whose fitness is determined by three variables: (i) the proportion of conspecifics in the population that the foundress can breed with; (ii) fitness of offspring of conspecific matings (between 0-1); and (iii) fitness of offspring of heterospecific matings (between 0-1). Further, we divide a syconium into five layers - at the centre, all eggs survive, while zero survive at the syconium edge (with 0.75, 0.5, 0.25 survival in the intermediate layers). We compare a '*CI*' and a '*no CI*' model across variation in our three variables. See Fig. 4 for a schematic outline of the parameters.

For example, for the *no CI* model at 75% conspecifics mating opportunities: 75% of eggs get conspecific relative mating fitness (e.g., $\omega = 0.8$) and 25% of eggs get heterospecific relative mating fitness (e.g., $\omega = 0.2$). Each egg gets randomly assigned into a layer of the syconium and each egg's relative mating fitness is multiplied by its oviposition site fitness (i.e., randomly assigned: 1.0, 0.75, 0.5, 0.25, 0). Finally, all 1000 eggs' scores are summed to obtain inclusive fitness for the foundress. For the *CI* model (at 75% conspecifics mating opportunities) we lose 25% of the egg load but all eggs get the conspecific relative mating fitness (i.e., 0.8) - then they are randomly placed into the best remaining 750 oviposition site positions (mimicking that fig wasps preferentially lay in optimum sites). Scores are again multiplied and summed across 10 replicates. The key modeling assumption here is that unviable eggs are not oviposited and thus do not waste premium oviposition sites (this is also equitable to a scenario where egg oviposition order is prioritised). Results are obtained by subtracting realised *CI* versus *no CI* fitness across all combinations of conspecific-heterospecific relative mating fitness (i.e., 0-1 for both) surfaces. We display individual heat maps for outcomes at different percentages of conspecific mating opportunities.

Results

Wolbachia screening of field collected samples

From 253 screened wasps, 41.1% (104 individuals) displayed *Wolbachia* infection. We found no evidence for multiple infections as no chromatograms contained multiple peaks. Individual *wsp* and MLST phylogenies confirm the monophyly of identified strains (Table S1; Fig. S1); the *wsp* tree contained five major clades while the MLST tree contained six. *wsp* clade assignment mostly matched with MLST markers, with the exception of the MLST clades associated with *F. trichocerasa* subsp. *pleioclada* and *F. microdyctia* , which were contained within the same *wsp*clade (Table S1; Fig. S1). We therefore followed *wsp* clade designation for all wasps apart from *wsp* clade six which was split into two (wspC6_1 and wspC6_2) giving six identified *Wolbachia* clades.

While 77% of all *Ceratosolen armipes* (the pollinator of lowland*F. itoana*) were infected with *Wolbachia*, only one (out of 34) of the mid-elevation *C.* sp (*ex.* mid-elevation*Ficus umbrae*) was infected. Similarly, ca. 63% of *C. "kaironkensis"* (*ex.* highland *F. microdyctia*) were infected. These wasps and their fig hosts form monophyletic clades with species replacing each other in parapatry (Fig. 3), mid-elevation *F. umbrae* (largely *Wolbachia* free) and highland *F. microdyctia* being sister species and lowland *F. itoana* the outgroup (Souto-Vilaros et al., 2018). Alternative populations of the four *Ceratosolen* pollinator species associated with the single species *F. arfakensis* showed disjunct infection statuses, with 26% infected overall but with different infection frequencies and strains at different elevations (e.g., *wspC3* in the lowlands and *wsp C2* in the highlands). For *F. trichocerasa* (a single species comprising two distinct subspecies) the proportion of infected wasp pollinators differed between host fig subspecies with lowland subsp. *trichocerasa* and highland subsp. *pleioclada* having 84% and 54% infection frequency, respectively. Strain identity was also largely distinct to a given subspecies. In the case of *F. wassa* (a genetically homogenous entity across the gradient hosting a two major pollinator clades), only 10% of pollinator wasps (all individuals from highland populations) were infected. Overall, sister species/populations of wasps usually had different *Wolbachia* infection status or strain type (Fig. 3 & Fig. S1). These sister species of wasps were not infected by monophyletic MLST or *wsp* (except *wspC6*) clades of *Wolbachia*.

wsp strains appear restricted to lowlands or highlands. For instance, *wsp* clades 1, 6_1 and 6_2 are present in wasps from elevations above 2,200m while the rest occur in the lowlands (below 1,200m). An exception is for wasps originating from the mid-elevation site (here considered as 1,700m) "Degenumbu" where both lowland and highland strains occur. For instance, *wsp* clade 1 (a highland strain) occurs in *F. wassa* wasps from this location; similarly, *wsp* clade 2, a lowland strain occurs in *F. arfakensis* wasps from 1,700m. Overall, bar a few exceptions, strain type segregates by (sub)species while infection status seems to be influenced by elevation.

Simulation of Wolbachia *distribution among host species under the 'contact contingency' hypothesis*

Our wolPredictor simulation was able to predict positive strain associations at up to 88.46% (92/104 individuals at species clustering levels of 10-13; SI runs '*pleio* ' 4, 5, 12 & 18) accuracy against the empirically observed infection statuses across our fig wasp phylogeny (Fig. 4). Predictive accuracy of greater than 80% was found in 16 of 20 runs at species clustering levels of 10-19. Investigation at these species delimitation assessments show high congruence with species diversity patterns in Souto-Vilarós et al. (2019), notably with wasps from *F. arfakensis* , *F. pleioclada* and *F. wassa* split into two or three putative species featuring alternate *Wolbachia* strain statuses. The highest accurate overall strain prediction, 65.61%, (30 positive and 136 negative predictions) regularly occurred at species clustering levels of 5-7 – with the *wolPurger* function removing around 30 positive predictions and adding >100 negative predictions. In general, improved negative strain accuracy often trades-off with losses in correct positive strain predictions. High non-infection prediction accuracy occurs at lower species clustering levels where large singleton clades within communities are ascribed negative *Wolbachia* associations.

One-sample T-tests of these best-scoring results against 1000 randomly generated predictions (mean = 14.16% accuracy for positive strains only; mean = 37.86% accuracy including negative strains) indicates that our model simulation predicts *Wolbachia* infection status with significantly higher accuracy (t = -

698.98, d.f. = 999, p < 2.2e-16 for positive strains only; t = -335.61, d.f. = 999, p < 2.2e-16 including negative strains). As a further control, we also ran 100 wolPredictor simulations (see SI files:*pleio_shuff**) with assayed *wsp* clades randomly shuffled – the best predictive power for positive strains fell to around 46.15% (mean = 40.19%) significantly less than our best prediction for positive strains (t = -33542, d.f. = 99, p < 2.2e-16).

Simulation of fig wasp inclusive fitness outcomes under differential hybrid embryo mortalities according to the 'oviposition trade-off' hypothesis

Inclusive fitness of individual wasp foundresses differed according to the imposition of CI-induced egg mortality (Fig. 5). As the population level of conspecific mating increases, a greater number of pixels (representing relative conspecific-heterospecifc mating fitness values) begin to favour the CI-induced egg mortality model (Table 1). Low levels of conspecific mating do not favour CI as minimal gains in favourable ovipositioning outcomes do not outweigh large-scale fecundity losses of heterospecific matings. At greater levels of conspecific mating (>60%), even marginal relative fitness differences between conspecific and heterospecific offspring (e.g., 0.55 versus 0.45, respectively) result in higher inclusive fitness for foundresses operating under CI-induced egg mortality. This is because favourable ovipositioning positively trades-off against reductions in foundress fecundity.

Discussion

Understanding the historical eco-evolutionary processes that determine the structure of biodiversity are primary goals in ecology and evolutionary biology. Therefore, the current understanding that a large proportion of arthropod diversity (Zug & Hammerstein 2015) harbours a non-systematically distributed agent of speciation constitutes a major academic challenge with respect to identifying predictable processes that generate biodiversity. Here we introduce the 'contact contingency' predictive model for *Wolbachia* strain distributions based on phylogenetic relationships, ecological contact, and host abilities to determine their own infection status, that shows remarkable accuracy on an empirical fig wasp dataset sampled from elevational transects in Papua New Guinea. We also present the 'oviposition trade-off' model to account for the influence of post-zygotic fitness losses imposed by *Wolbachia* which could potentially invalidate our proposal. We show that post-zygotic fecundity losses may be tolerated under multiple-mating scenarios when considering the ecological characteristics of the fig syconia microcosm. We hope our work stimulates further debate around these phenomena in order to either accept and refine our proposed mechanism, or to prompt a dismissal accompanied by a more parsimonious theoretical framework that accounts for distributions of *Wolbachia* endosymbionts at the host taxonomic scales we are querying. Our models may be particularly suited to the unusual ecology of fig wasps, but their underlying dynamics may contribute to differing degrees among other ecological systems (Box 1).

An inspection of *Wolbachia* strain distributions among our fig wasp phylogenies reveals suggests that systemic processes may be in operation. Wasp populations at different elevations display different infection patterns. For example, 77% of *Ceratosolen armipes* are infected but only one out of 34 individuals of its clade that pollinates *Ficus umbrae* carries *Wolbachia* . Further, there are *Wolbachia* strains that are restricted to lowland and highland elevations, even among the same host (e.g., *F. arfakensis*). As such, parapatric populations of pollinating wasps often have contrasting infection status that may represent RI, mediated by either uni- or bi-directional CI. Our data repudiate *vertical transmission co-divergence* hypotheses. However, we consider it important to note that this also repudiates a *horizontal exchange hypothesis*, as *Wolbachia* strains invariably do not occupy all or multiple wasp clades infecting the same host species/complex, despite potential for ecological contact. These results are consistent with the idea that *Wolbachia* only infects certain insect groups under particular host-adaptive circumstances.

Systemically generated patterns in host *Wolbachia* statuses could arise if infection is largely controlled by endosymbiont abilities to manipulate their own infection status as is often posited (Werren *et al.* 2008); for example, because particular clades of host lineages' immune responses are unable to repel the infection. However, our observed patterns do not obviously support manipulation by the endosymbiont given the

7

correlation with abiotic drivers such as altitude, and no known *a priori* reason to assume that host immunity is principally determined by abiotic circumstances. Relative fitness costs in host mating potential, fecundity, or inappropriate sex ratio distortion associated with *Wolbachia* invasion may also generate the conclusion that *Wolbachia* is the chief architect of its own success. However, it is known that *Wolbachia* offers host fitness benefits that may have ecological contingencies (Zug & Hammerstein 2015; Correa & Ballard 2016). We may consider a host insect species that exhibits a broad phenotypic range, say for ovipositor length that has distinct optima at different altitudinal ranges according to host plant morphological divergence. Any mechanism that prevented reproductive events between extreme phenotypes (yielding intermediate morphs) at inappropriate altitudes would be favoured providing the benefits (increased local fitness) outweighed any deleterious costs.

Our wolPredictor simulation of the 'contact contingency' hypothesis attempts to model such a scenario that is founded on: (i) host abilities to modulate/purge *Wolbachia* when evolutionarily apposite; and (ii) *Wolbachia* causing a circumstantial host fitness advantage. Our hypothesis is that recently evolved sister-species diversifying within the same host community are at a selective advantage when they harbour alternate strains of *Wolbachia*. Divergent strains facilitate the initial stages of speciation; this is because we expect that closely related, co-diverging *Wolbachia* strains would not confer RI between hosts. Given that *Wolbachia* infection does appear to contain some fitness cost component (Zug & Hammerstein 2015), we predict that these patterns should not be evident among diverging lineages that are not in regular ecological contact, thus leaving *Wolbachia* subject to host immunosuppression due to the circumstantial absence of any RI-enforcing benefit. Finally, as *Wolbachia* typically drops out of host lineages after approximately 7 million years (Bailly-Bechet *et al.* 2017), and because alternative mechanisms of RI that require cytogenetic or morphological modification may take longer to evolve (Bordenstein *et al.* 2001), we predict that closely-related species that have not recently diverged should also be purged of *Wolbachia* infection. This would reflect a hypothesis that observed lineage dropout (Bailly-Bechet *et al.* 2017) results from temporal changes in the adaptive benefits of *Wolbachia* , which may subsequently become redundant (the *adaptive decay* hypothesis).

Simulation results among our fig wasp data are particularly impressive (ca. 88% accuracy for positive strains) when the species delimitation algorithm matches the empirical understanding of ecological diversity patterns (Souto-Vilarós *et al.* 2019). In particular, when multiple congeners within fig host communities are algorithmically predicted among *F. arfakensis* , *F. trichocerasa* subsp. *pleioclada* and *F. wassa* , wolPredictor ascribes multiple strains within these communities that reflect the data. Prediction accuracy when also considering negative (uninfected) *Wolbachia* strains is less precise. This may be due to variation in the empirical data where negative and positive strains co-occur within clades or the tendency for non-infected individuals to appear among lineages that comprise multiple/incipient species within a single host fig (e.g., *F. wassa* - infection rate = ca. 10%), where wolPredictor ascribes positive *Wolbachia* associations (such patterns may actually represent unidirectional RI whilst we model the bidirectional mechanism).

To model the 'adaptive decay' hypothesis, our *wolPurger* function operates to remove *Wolbachia* from lineages after extended evolutionary timescales. However, as it uniformly removes *Wolbachia* across all samples exceeding the distance threshold it constitutes a crude method that may not always be appropriately applied across lineages. Inconsistencies in strain associations could result from other drivers we have not accounted for, and alternative mechanisms of RI (hypothetically rendering *Wolbachia* redundant) may appear at different rates across lineages possibly due to serendipitous genomic architecture or unconsidered ecological contingencies. For example, among fig wasps, syconia access is partially controlled by relative syconia-wasp size (Bronstein 1987), which mechanically prevents hybridisation opportunity among some species.

Furthermore, under a simple model of panmixis and infinite population size, CI is predicted to sweep to fixation, contrary to the population level polymorphism observed in our data. Theoretical expectations on how CI spreads through populations are largely determined by population structure (Engelstädter & Telschow 2009). However, this depends on perfect transmission and infection rates may still eventually decay even if fixation is achieved. Fig wasps are both haplodiploid and inbred. Haplodiploidy can facilitate the survival of infected haploid males (Breeuwer & Werren 1990), which like inbreeding can result in both a

8

higher invasion threshold and a reduced stable equilibrium frequency (Engelstädter & Hurst 2006). These considerations deserve further attention, but they may, at least in part, explain infection frequencies below 100% as observed in our study system.

Overall, our wolPredictor simulation of the 'contact contingency' hypothesis is a rule-based algorithm that manages to capture much of the embedded structure in a dataset that presents a superficially stochastic appearance. Thus, it suggests that some environmentally contingent symbiotic benefits (Correa & Ballard 2016) may systematically sum to yield predictable *Wolbachia* distributions. Our methods cannot test for precise mechanisms determining *Wolbachia* distributions across our study systems, and the algorithm underpinning wolPredictor may inadvertently represent some other set of real-world contingencies.

Moreover, we acknowledge that our model may be considered problematic since *Wolbachia* -mediated CI is a post-zygotic mechanism that elicits an immediate fitness cost in host fecundity. However, it is feasible that the unique life-histories and ecological conditions of fig wasps means they may be tolerant of CI: oviposition sites are at especially high premium (Dunn *et al.* 2015), fig wasps are known to produce surplus eggs (Dunn *et al.* 2011), and co-evolved species are renowned for precise tolerances in interacting traits that may render hybridisation particularly costly. This critique prompts our 'oviposition trade-off' hypothesis and second simulation model. Here we investigated the impact of CI in multiple-mating scenarios when considering the oviposition limiting constraints of fig syconia. We show that inclusive fitness of multiple mated females can be higher despite fecundity losses providing that egg load reduction (or selective ovipositioning) facilitates oviposition into higher-quality fig ovules that are less vulnerable to parasitoid attack (Dunn *et al.* 2008). Notably, the parameter values (i.e., conspecific mating levels and relative fitnesses) that yielded CI favouring outcomes are realistically achievable among natural fig wasp populations. The results imply that bi-directional CI may adaptively evolve in fig wasps without accompanying mechanisms.

The interaction of CI-inducing *Wolbachia* on multiple-mating in fig wasps has not been studied but we found a single study in *Drosphila* demonstrating *Wolbachia* associated reductions in sperm competition abilities (Champion de Crespigny & Wedell 2006). Given the unusual reproductive manipulations of haplodiploid Hymenoptera such as selective fertilisation, adjustment of sex-ratios, and control over oviposition order according to ploidy, and given the dearth of research of CI under multiple-mating conditions, it is entirely possible that such dynamics are at play among fig wasps at least. Thus, future work examining whether incompatible matings result in differential pre-oviposition embryo mortality, and whether selective oviposition of conspecific versus heterospecific eggs occurs would be of great value. It may be that RI-inducing *Wolbachia* constitutes a mutualist symbiont preventing the inefficient use of highly valuable oviposition sites with intermediate hybrid form, lower fitness offspring.

We further note that our models diverge from some conventionally held opinions regarding *Wolbachia* and host manipulation/purging (e.g., Werren *et al.* 2008) capabilities and a general view that *Wolbachia* is in conflict with its hosts in many respects (e.g., Charlat *et al.* 2007). While it has been shown that *Wolbachia* may employ microRNAs to alter host gene expression (Hussain *et al.* 2011), it has not been investigated whether any Red Queen-type arms race dynamics facilitate host resistance – as our model implies. We argue that despite potential pitfalls it is difficult to propose an alternative systemic model or explain observed structural patterns as random, within our or other published datasets. For example, an overview of malvantheran fig wasps shows that communities featuring singleton congeners invariably display negative *Wolbachia* associations while the reverse is true for multi-congener communities (Haine & Cook 2005). Finally, we also contend that the most parsimonious interpretation of reported empirical patterns (i.e., sister-species hosting paraphyletic infections) leads to a view that the interests of *Wolbachia* may be aligned with their hosts under conditions of ecological speciation. *Wolbachia* is known to impart some host benefits (Zug & Hammerstein 2015) and we are yet to fully understand the nuances, trade-offs and ecological contingencies that determine whether it is rendered circumstantially advantageous.

Our proposed models are particularly suited to testing in fig wasps due to the high degree of easily-collatable host-specificity relationships exhibited when wasps co-occur within the fig microcosm. Such processes may also be subtly at play among other taxa where ecological contact is not easily assessed – such data have not

9

been previously incorporated into a predictive phylogenetic model. Variables such as ecological contact may act along a continuum among species rather than the easily assessable binary states observed in fig wasps. Indeed, a suite of variables that comes into play in fig wasp communities, e.g., degrees of inbreeding or variation in oviposition site quality, may differentially contribute to different systems and may interact with other factors regarding *Wolbachia* cost-benefit trade-offs that determine infection status. Thus, fig wasps may offer an ideal window into understanding not only the determinant ecological contingencies at play but also offer insight into the nature of the types of variables that may be significant or even overridden by other factors (for example, we note that our 'contact contingency' hypothesis is pertinent to all fig wasps whereas the 'oviposition trade-off' hypothesis is relevant to pollinators and non-pollinators that enter the fig). Even with these system-specific limitations in mind, the vital issue then may be that for most of global diversity we simply do not have the kind of detailed ecological information that can reliably inform us about the processes underpinning community assembly. Our theoretical reasoning may therefore be extremely generalizable and we urge that model communities representing different ecosystem states are identified and investigated (Box 1) under a proposed methodological framework (Box 2).

Our models imply a genomic mechanism among insects making *Wolbachia* tolerance/purging a highly labile and evolutionarily unstable trait that would concomitantly render RI events a relatively trivial occurrence. Other potentially similar mechanisms have been documented in other taxa (see also Box 1). In the hymenopteran *Nasonia*, *Wolbachia* -induced CI has been shown to precede other incompatibility mechanisms (Bordenstein *et al.* 2001). In *Drosophila* , readily occurring genomic inversions (Noor & Bennett 2009) often serve to maintain RI among closely-related species/subspecies in sympatry (although CI is also found in fruit-flies; Merçot & Charlat 2004), while such phenomena are less prominent in allopatry (Noor *et al.* 2001). And among certain gastropods, a single gene mutation coding for shell chirality can cause RI between sister-species (Hoso *et al.* 2010). Under the *islands of speciation* paradigm (Noor & Bennett 2009), regions of genomic divergence incrementally build until differences between lineages yield distinct evolutionary trajectories. A conundrum exists in explaining how this might generate entirely separate lineages under ecologically driven divergence in sympatry and in the face of gene flow. Future work may reveal that it is highly labile, binary decision-making reproductive switches such as CI or chromosomal inversions that resolve this puzzle by providing tipping point mechanisms that promote evolutionary schisms when net selective pressures favour speciation.

The stark disparity between recent advances in genomic data accrual relative to the laborious efforts required to record phenotypic/ecological data is well acknowledged – our work highlights the need to ameliorate this. Biodiversity cannot be simplistically evaluated as a metric indicating number of species nor considered solely as the outcome of interactions between closely-related species within the same trophic level. There is a growing consensus that we need to consider interactions both within and between all trophic levels whilst also identifying what constitutes significant versus trivial dynamics (i.e., intensity of interaction; Segar *et al.* 2020), or, more generally, ecological contingency. Thus, we face a massive challenge to document community dynamics of not only obviously tractable relationships (e.g., competition between focal species) but also of both mutualistic (e.g., pollination, seed dispersal, nutrient sequestration, mycorrhizal) and antagonistic factors (e.g., parasitic, disease) that may sometimes be bacterial, fungal or viral in origin. Failure to account for such agents may mean we never fully disentangle the myriad determinants of ecosystem dynamics nor quantify the relative contributions of stochastic (*viz* . neutral; Hubbel 2001) processes.

CONCLUSIONS Our results indicate that *Wolbachia* distributions are predictably structured among an arthropod dataset based on a predictive model invoking adaptive responses in host fig wasps. A parsimonious interpretation of these findings suggests that ecologically contingent co-evolutionary benefits of *Wolbachia* induced CI, particularly with respect to opportunity for lineage diversification, may systematically sum to yield predictable distributions despite initial appearances that the endosymbiont is stochastically distributed at some taxonomic resolutions. In particular, our data suggests that future work assessing biodiversity patterns among arthropods should incorporate *Wolbachia* infection data (alongside other microorganisms) as an added modelling dimension in order to account for a potentially confounding variable. Our aim is to stimulate debate and subsequent research in unravelling a rather puzzling phenomenon within arthropod

biodiversity.

TABLE LEGENDS

Table 1. Table indicating percentage of pixels where *CI* model is favoured over *non-CI* according to level of conspecific mating.

Table S1. A table of closest match alleles in the MLST data base for each strain and gene. Taxa are selected to be representative of their strain. Numbers correspond to allele number in the MLST data base and phylogenies above (e.g. SnoWas27 is a representative of the wspC1 clade, the allele that is closest in the MLST data base for its hcpa sequence is 80). Note that wspC6_1 and wspC6_2 share the same nearest match allele for wsp.

FIGURE LEGENDS

Figure 1. Conceptual diagram outlining the 'contact contingency' hypothesis. Hypothetical fig wasp relationships and predicted status of RI inducing *Wolbachia* according to variation in ecological contact and evolutionary time since speciation. We predict *Wolbachia* infection to occur only in community III where species 1 & 2 should harbour unrelated strains. Sister species 3 & 4 are not in ecological contact as they form separate communities I and II, while sister-species 5 & 6 in community IV, despite ecological contact, have had sufficient evolutionary time for alternative (less costly) RI mechanisms to evolve.

Figure 2. Stylised schematic showing a fig in cross section. Five layers of ovules are used in our model (white and grey) and no oviposition occurs in the central lumen (black). Ovule length (and embryo relative fitness, ω) decreases towards the fig wall (green) where larvae are at greater risk of parasitism. We use a descriptive model to contrast inclusive fitness (W) between foundress wasps that do not experience cytoplasmic incompatibility (wasp 1, blue) and those that do (wasp 2, orange). Here, in a toy example, each foundress has 10 eggs (open circles represent viable hybrid eggs with decreased fitness while closed circles are non-hybrids with full fitness) and we limit oviposition to two eggs per layer. While CI wasps lay fewer eggs (hybrids are lost to CI) they do not fill valuable oviposition sites with hybrids of decreased fitness. Here, the CI wasp gets an inclusive fitness of 3.8 for its seven remaining eggs and the noninfected wasp gets 3.1 for a full complement of 10 eggs (i.e., by multiplying egg fitness by oviposition fitness then summing). Inclusive fitness is therefore greater in wasp 2 despite this fecundity loss, as it lays a higher number of high fitness eggs in premium oviposition sites. This example would represent one pixel on the heat maps displayed in Figure 5. Please see text for further details.

Figure 3. *Wolbachia* strains mapped along the pollinating wasp phylogeny. Strain type is indicated by the different colours, with uninfected individuals in black. For each wasp clade the *Ficus*host is given.

Figure 4. *Wolbachia* strain prediction accuracy by wolPredictor modelling the 'contact contingency' hypothesis across 253 fig wasp pollinator samples at species clustering thresholds of 2-50 for run name *pleio\sout4* . Accurate positive assignations (orange) are shown above the zero-line whilst accurate negative assignations (blue) are shown as positive values below the zero-line.

Figure 5. Heat map evaluation of the 'oviposition trade-off' hypothesis. Comparative inclusive fitness values of fig wasp foundresses across relative conspecific-heterospecific fitness space at different population-level frequencies of conspecific mating (between 5-95%) under alternate scenarios of CI-induced egg mortality (i.e., '*CI* ' vs. '*no CI* '). Redder tones (i.e., above zero) indicate relative conspecific-heterospecific fitness where foundress inclusive fitness is higher under CI-induced mortality due to preferential oviposition of higher

fitness conspecific offspring despite trade-offs with fecundity reduction. NB in order to explore all relative fitness space, heatmaps indicate regions where heterospecific fitness is greater than conspecific fitness, which will generally be an unrealistic scenario.

Figure S1. Single phylogenies for (a) *coxA* , (b) *fbpA* , (c)*ftsZ* , d) *gatB* , (e) *hcpa* and (f) *wsp*including sequences generated for this study and all accessions from the MLST data base, (g) a *wsp* single gene phylogeny for sequences generated in this study and (h) a phylogeny derived from the five MLST genes for sequences from this study only. Colour coding depicts *Wolbachia* strains red: wspC1, blue: wspC2, purple: wspC3, green wspC5, yellow: wspC6_1 and orange: wsp6_2.

TABLE

Table 1.

% conspecifics	CI favoured (%)
5	2.18
10	5.16
15	8.46
20	12.12
25	14.63
30	17.52
35	20.89
40	24.78
45	27.05
50	29.78
55	33.08
60	37.32
65	39.2
70	41.47
75	44.88
80	49.85
85	49.85
90	49.94
95	49.34

FIGURES

Figure 1.
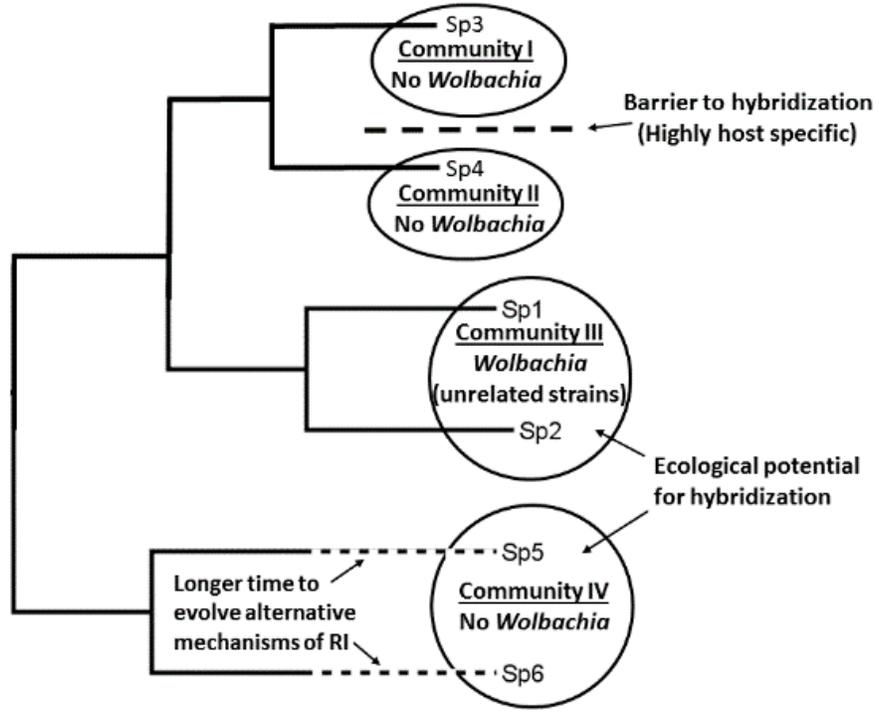
Figure 2.
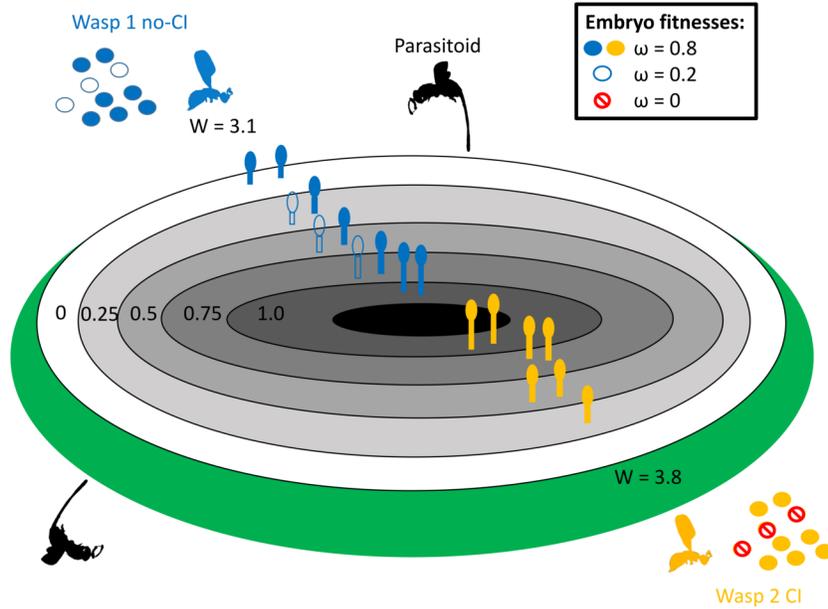
Figure 1: This is a caption

Figure 3.
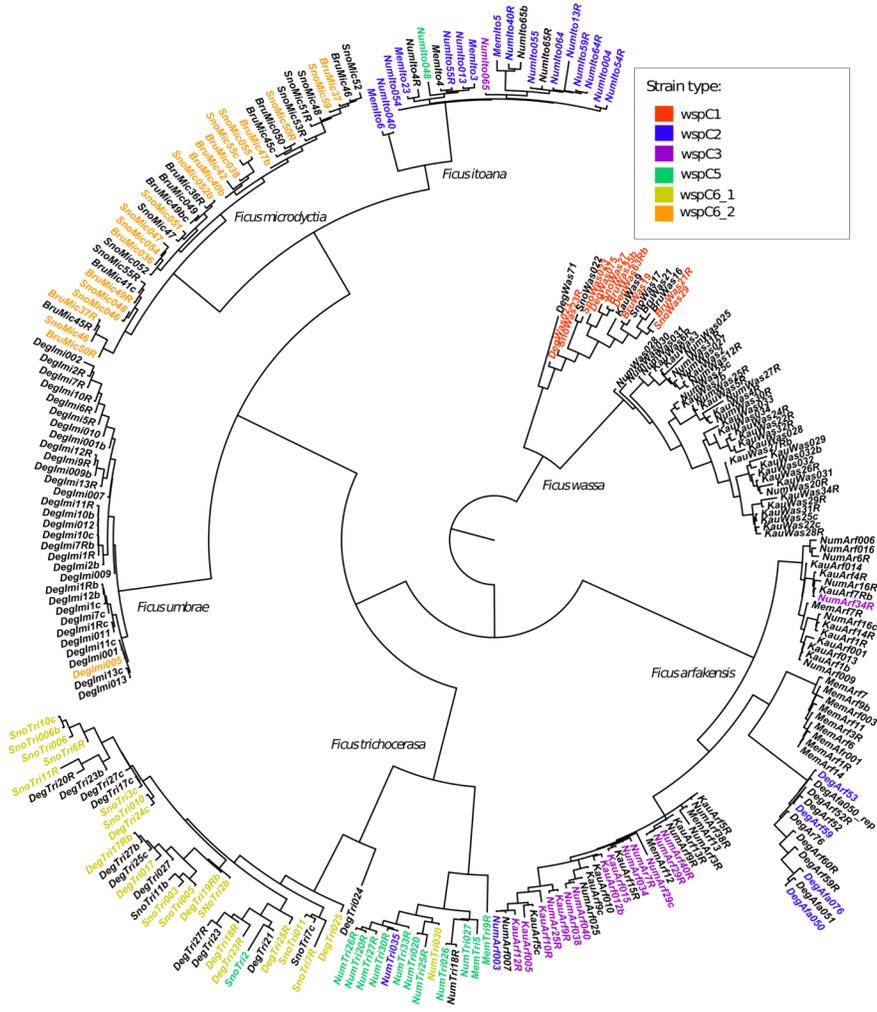

Figure 2: This is a caption

Figure 4.
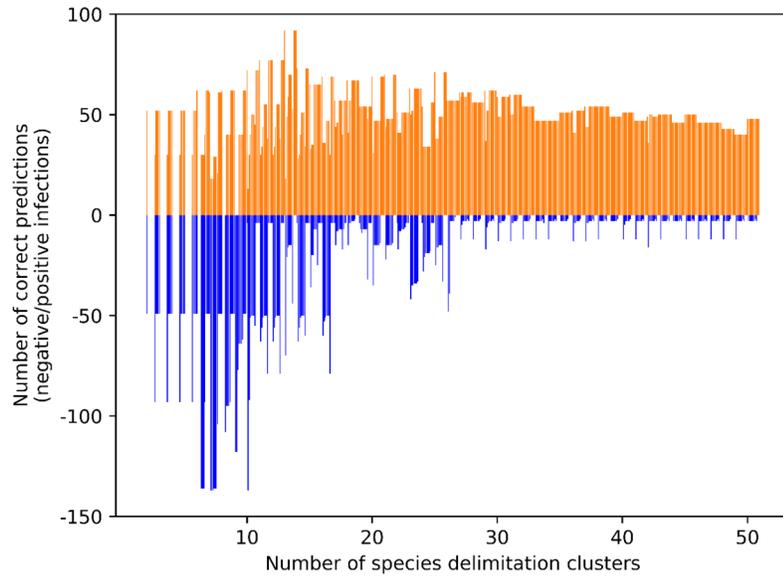
Figure 5.
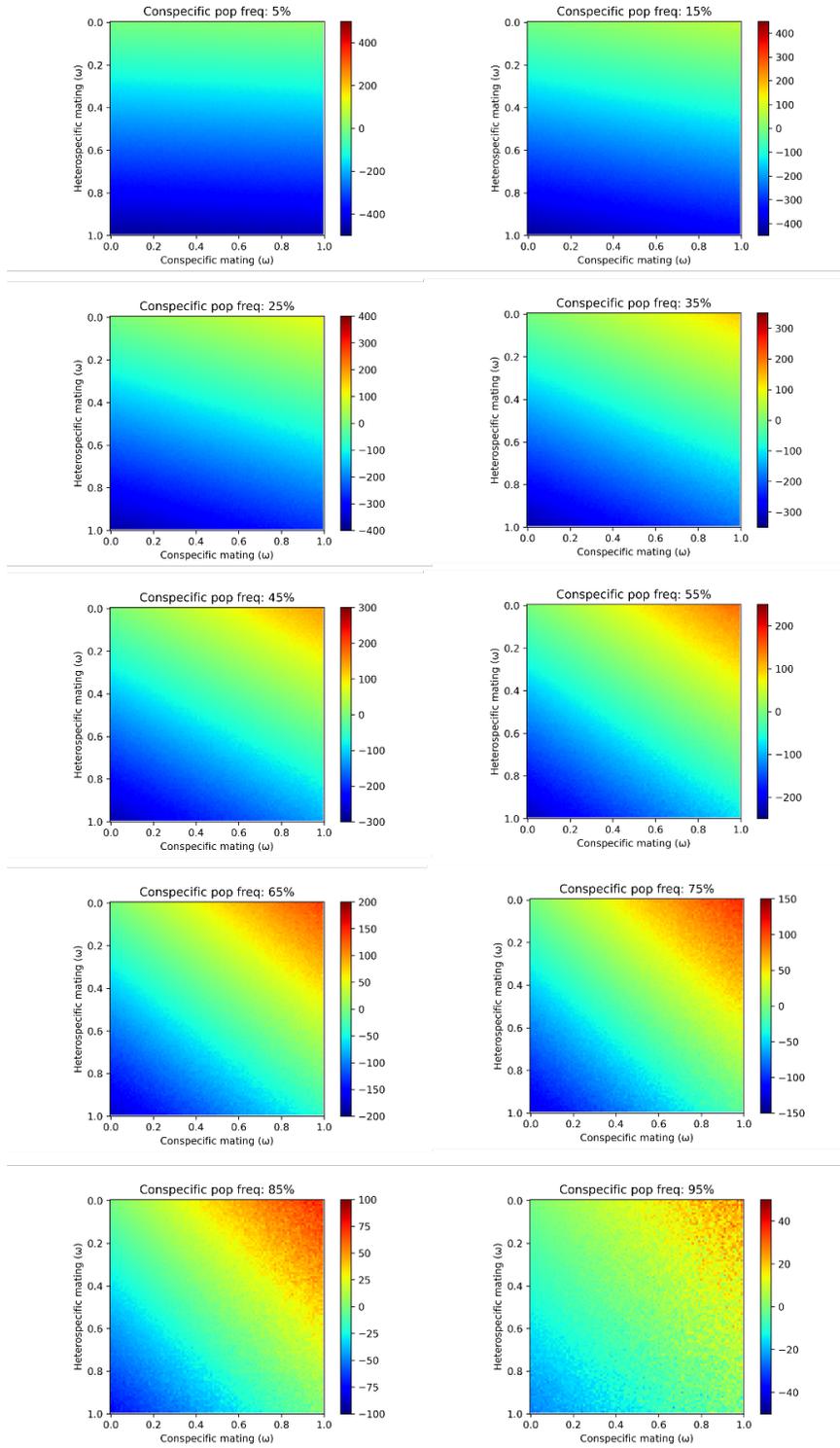
BOXES

Box 1.

Box 1. Beyond fig wasps: seeking generality across systems. Our insights derive from an unusual system comprisin

Box 2.

Box 2. A comparative approach to further derive process from phylogenetic patterns of *Wolbachia* distribu

References

Aberer AJ, Kobert K, Stamatakis A (2014) ExaBayes: massively parallel bayesian tree infer-ence for the whole-genome era.*Molecular Biology and Evolution*, **31, 2553–2556.**

Althoff DM, Segraves KA, Smith CI, Leebens-Mack J, Pellmyr O (2012) Geographic isolation trumps coevolution as a driver of yucca and yucca moth diversification. *Molecular Phyloge-netics and Evolution*, **62, 898–906.**

Bailly-Bechet M, Martins-Simões P, Szöllosi GJ *et al.*(2017) How long does wolbachia remain on board? *Molecular Biology and Evolution*, **34, 1183–1193.**

Baldo L, Dunning Hotopp JC, Jolley KA *et al.* (2006) Multilocus sequence typing system for the endosymbiont Wolbachia pipientis. *Applied and Environmental Microbiology*, **72, 7098–7110.**

Balvín O, Roth S, Talbot B, Reinhardt K (2018) Co-speciation in bedbug Wolbachia parallel the pattern in nematode hosts.*Scientific Reports*, **8, 8797.**

Beckmann JF, Bonneau M, Chen H *et al.* (2019) The Toxin-Antidote Model of Cytoplasmic Incompatibility: Genetics and Evolutionary Implications. *Trends in Genetics*, **35, 175–185.**

Beckmann JF, Ronau JA, Hochstrasser M (2017) A Wolbachia deubiquitylating enzyme in-duces cytoplasmic incompatibility.*Nature Microbiology*, **2, 17007.**

Bordenstein SR, Bordenstein SR (2016) Eukaryotic association module in phage WO genomes from Wolbachia. *Nature Communications*, **7, 13155.**

Bordenstein SR, O'Hara FP, Werren JH (2001) Wolbachia-induced incompatibility precedes other hybrid incompatibilities in Nasonia.*Nature*, **409, 707–710.**

Bordenstein SR, Werren JH (2007) Bidirectional incompatibility among divergent Wolbachia and incompatibility level differences among closely related Wolbachia in Nasonia. *Heredity*, **99, 278–287.**

Branca A, Vavre F, Silvain J-F, Dupas S (2009) Maintenance of adaptive differentiation by Wolbachia induced bidirectional cytoplasmic incompatibility: the importance of sib-mating and genetic systems.*BMC Evolutionary Biology*, **9, 185.**

Breeuwer JA, Werren JH (1990) Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature*, **346, 558–560.**

Bronstein JL (1987) Maintenance of Species-Specificity in a Neotropical Fig: Pollinator Wasp Mutualism. *Oikos*, **48, 39.**

Bronstein JL, Wilson WG, Morris WF (2003) Ecological dynamics of mutualist/antagonist communities. *The American Naturalist*, **162, S24–39.**

17

Casiraghi M, Anderson TJ, Bandi C, Bazzocchi C, Genchi C (2001) A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of Wolbachia endosymbionts. *Parasitology*, 122 Pt 1, 93–103.

Champion de Crespigny FE, Wedell N (2006) Wolbachia infection reduces sperm competitive ability in an insect. *Proceedings. Biological Sciences / the Royal Society*, 273, 1455–1458.

Charlat S, Hornett EA, Fullard JH *et al.* (2007) Extraordinary flux in sex ratio. *Science*, 317, 214.

Cook JM, Segar ST (2010) Speciation in fig wasps.*Ecological Entomology*, 35, 54–66.

Correa CC, Ballard JWO (2016) Wolbachia Associations with Insects: Winning or Losing Against a Master Manipulator. *Frontiers in ecology and evolution*, 3.

Coyne JA, Orr HA (2004) *Speciation.* Sinauer Associates is an imprint of Oxford University Press, Sunderland, Mass.

Crespi BJ, Morris DC, Mound LA (2004) *Evolution of ecological and behavioural diversity: Australian Acacia thrips as model organisms.* cabdirect.org.

Darwell CT, al-Beidh S, Cook JM (2014) Molecular species delimitation of a symbiotic fig-pollinating wasp species complex reveals extreme deviation from reciprocal partner specificity. *BMC Evolutionary Biology*, 14, 189.

Darwell CT, Rivers DM, Althoff DM (2016) RAD-seq phylogenomics recovers a well-resolved phylogeny of a rapid radiation of mutualistic and antagonistic yucca moths. *Systematic entomology*, 41, 672–682.

Dedeine F, Vavre F, Fleury F *et al.* (2001) Removing symbiotic Wolbachia bacteria specifically inhibits oogenesis in a parasitic wasp. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 6247–6252.

Dunn DW, Jander KC, Lamas AG, Pereira RAS (2015) Mortal combat and competition for oviposition sites in female pollinating fig wasps.*Behavioral Ecology*, 26, 262–268.

Dunn DW, Jansen-González S, Cook JM, Yu DW, Pereira RAS (2011) Measuring the discrepancy between fecundity and lifetime reproductive success in a pollinating fig wasp. *Entomologia Experimentalis et Applicata*, 140, 218–225.

Dunn DW, Segar ST, Ridley J *et al.* (2008) A role for parasites in stabilising the fig-pollinator mutualism. *PLoS Biology*, 6, e59.

Engelstädter J, Hurst GDD (2006) Can maternally transmitted endosymbionts facilitate the evolution of haplodiploidy? *Journal of Evolutionary Biology*, 19, 194–202.

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

Gavrilets S (2004) *Fitness landscapes and the origin of species (MPB-41).* books.google.com.

Gerth M, Röthe J, Bleidorn C (2013) Tracing horizontal Wolbachia movements among bees (Anthophila): a combined approach using multilocus sequence typing data and host phylogeny. *Molecular Ecology*, 22, 6149–6162.

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, 27, 221–224.

Greeff JM, van Noort S, Rasplus J-Y, Kjellberg F (2003) Dispersal and fighting in male pollinating fig wasps. *Comptes Rendus Biologies*, 326, 121–130.

Haine ER, Cook JM (2005) Convergent incidences of Wolbachia infection in fig wasp communities from two continents.*Proceedings. Biological Sciences / the Royal Society*, 272, 421–429.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.*Nucleic acids symposium series.*

Hamilton WD (1967) Extraordinary sex ratios. A sex-ratio theory for sex linkage and inbreeding has new implications in cytogenetics and entomology. *Science*, 156, 477–488.

Hamm CA, Begun DJ, Vo A *et al.* (2014) Wolbachia do not live by reproductive manipulation alone: infection polymorphism in Drosophila suzukii and D. subpulchrella. *Molecular Ecology*, 23, 4871–4885.

Hansen AK, Moran NA (2014) The impact of microbial symbionts on host plant utilization by herbivorous insects. *Molecular Ecology*, 23, 1473–1496.

Harmon LJ, Andreazzi CS, Débarre F *et al.* (2019) Detecting the macroevolutionary signal of species interactions.*Journal of Evolutionary Biology*, 32, 769–782.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings. Biological Sciences / the Royal Society*, 270, 313–321.

Hoso M, Kameda Y, Wu S-P *et al.* (2010) A speciation gene for left-right reversal in snails results in anti-predator adaptation.*Nature Communications*, 1, 133.

Hubbel SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton. 390 p.

Hussain M, Frentiu FD, Moreira LA, O'Neill SL, Asgari S (2011) Wolbachia uses host microRNAs to manipulate host gene expression and facilitate colonization of the dengue vector Aedes aegypti.*Proceedings of the National Academy of Sciences of the United States of America*, 108, 9250–9255.

Jäckel R, Mora D, Dobler S (2013) Evidence for selective sweeps by Wolbachia infections: phylogeny of Altica leaf beetles and their reproductive parasites. *Molecular Ecology*, 22, 4241–4255.

LePage DP, Metcalf JA, Bordenstein SR *et al.* (2017) Prophage WO genes recapitulate and enhance Wolbachia-induced cytoplasmic incompatibility. *Nature*, 543, 243–247.

McFrederick QS, Rehan SM (2016) Characterization of pollen and bacterial community composition in brood provisions of a small carpenter bee. *Molecular Ecology*, 25, 2302–2311.

Merçot H, Charlat S (2004) Wolbachia infections in Drosophila melanogaster and D. simulans: polymorphism and levels of cytoplasmic incompatibility. In: *Drosophila melanogaster, Drosophila simulans: So Similar, So Different* Contemporary issues in genetics and evolution. (eds Capy P, Gibert P, Boussy I), pp. 51–59. Springer Netherlands, Dordrecht.

Miller MA, Schwartz T, Pickett BE *et al.* (2015) A restful API for access to phylogenetic tools via the CIPRES science gateway.*Evolutionary Bioinformatics Online*, 11, 43–48.

Miraldo A, Duplouy A (2019) High wolbachia strain diversity in a clade of dung beetles endemic to madagascar. *Frontiers in ecology and evolution*, 7.

Moe AM, Weiblen GD (2012) Pollinator-mediated reproductive isolation among dioecious fig species (Ficus, Moraceae).*Evolution*, 66, 3710–3721.

Molbo D, Machado CA, Sevenster JG, Keller L, Herre EA (2003) Cryptic species of fig-pollinating wasps: implications for the evolution of the fig-wasp mutualism, sex allocation, and precision of adaptation.*Proceedings of the National Academy of Sciences of the United States of America*, 100, 5867–5872.

Murray MG (1990) Comparative morphology and mate competition of flightless male fig wasps. *Animal Behaviour*, 39, 434–443.

Noor MAF, Bennett SM (2009) Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, 103, 439–444.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 12084–12088.

Nosil P (2012) *Ecological Speciation*. Oxford University Press, Oxford, UK.

Novikov A (2019) Pyclustering: data mining library. *The Journal of Open Source Software*, 4, 1230.

Novotny V, Miller SE, Baje L *et al.* (2010) Guild-specific patterns of species richness and host specialization in plant-herbivore food webs from a tropical forest. *The Journal of Animal Ecology*, 79, 1193–1203.

Raja S, Suleman N, Compton SG, Moore JC (2008) The mechanism of sex ratio adjustment in a pollinating fig wasp. *Proceedings. Biological Sciences / the Royal Society*, 275, 1603–1610.

Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH (2009) Modes of acquisition of Wolbachia: horizontal transfer, hybrid introgression, and codivergence in the Nasonia species complex. *Evolution*, 63, 165–183.

Sakai S (2002) A review of brood-site pollination mutualism: plants providing breeding sites for their pollinators. *Journal of Plant Research*, 115, 161–168.

Segar ST, Fayle TM, Srivastava DS *et al.* (2020) The role of evolution in shaping ecological networks. *Trends in Ecology & Evolution*, 35, 454–466.

Segar ST, Volf M, Zima Jnr J *et al.* (2017) Speciation in a keystone plant genus is driven by elevation: a case study in New Guinean Ficus. *Journal of Evolutionary Biology*, 30, 512–523.

Shoemaker DD, Katju V, Jaenike J (1999) Wolbachia and the evolution of reproductive isolation between drosophila recens and drosophila subquinaria. *Evolution*, 53, 1157–1164.

Shoemaker DD, Machado CA, Molbo D *et al.* (2002) The distribution of Wolbachia in fig wasps: correlations with host phylogeny, ecology and population structure. *Proceedings. Biological Sciences / the Royal Society*, 269, 2257–2267.

Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR (2018) One prophage WO gene rescues cytoplasmic incompatibility in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 4987–4991.

Sintupachee S, Milne JR, Poonchaisri S, Baimai V, Kittayapong P (2006) Closely related Wolbachia strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant.*Microbial Ecology*, 51, 294–301.

Smith MA, Bertrand C, Crosby K *et al.* (2012) Wolbachia and DNA barcoding insects: patterns, potential, and problems. *Plos One*, 7, e36514.

Sneath PHA, Sokal RR (1973) *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. Freeman, San Fransisco.

Souto-Vilarós D, Machac A, Michalek J *et al.* (2019) Faster speciation of fig wasps than their host figs leads to decoupled speciation dynamics: snapshots across the speciation continuum.*Molecular Ecology*.

Souto-Vilarós D, Proffit M, Buatois B *et al.* (2018) Pollination along an elevational gradient mediated both by floral scent and pollinator compatibility in the fig and fig-wasp mutualism.*The Journal of Ecology*, 106, 2256–2273.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.

Telschow A, Flor M, Kobayashi Y, Hammerstein P, Werren JH (2007) Wolbachia-induced unidirectional cytoplasmic incompatibility and speciation: mainland-island model. *Plos One*, 2, e701.

Via S, Hawthorne DJ (2002) The genetic architecture of ecological specialization: correlated gene effects on host use and habitat choice in pea aphids. *The American Naturalist*, 159 Suppl 3, S76–88.

Weinert LA, Araujo-Jnr EV, Ahmed MZ, Welch JJ (2015) The incidence of bacterial endosymbionts in terrestrial arthropods.*Proceedings. Biological Sciences / the Royal Society*, 282, 20150249.

Wellenreuther M, Bernatchez L (2018) Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology & Evolution*, 33, 427–440.

Werren JH, Baldo L, Clark ME (2008) Wolbachia: master manipulators of invertebrate biology. *Nature Reviews. Microbiology*, 6, 741–751.

Yang C-Y, Xiao J-H, Niu L-M *et al.* (2012) Chaos of Wolbachia sequences inside the compact fig syconia of Ficus benjamina (Ficus: moraceae). *Plos One*, 7, e48882.

Yu H, Tian E, Zheng L *et al.* (2019) Multiple parapatric pollinators have radiated across a continental fig tree displaying clinal genetic variation. *Molecular Ecology*, 28, 2391–2405.

Zimmer C (2001) Wolbachia. A tale of sex and survival.*Science*, 292, 1093–1095.

Zug R, Hammerstein P (2015) Bad guys turned nice? A critical assessment of Wolbachia mutualisms in arthropod hosts. *Biological Reviews of the Cambridge Philosophical Society*, 90, 89–111.