A single synonymous nucleotide change impacts the male-killing phenotype of prophage WO gene wmk

Jessamyn I Perlmutter1,2,3*, Jane E Meyers1,3, Seth R Bordenstein1,3,4,5*

1Department of Biological Sciences, Vanderbilt University, Nashville, United States; 2Department of Molecular Biosciences, University of Kansas, Lawrence, United States; 3Vanderbilt Microbiome Innovation Center, Vanderbilt University, Nashville, United States; 4Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, United States; 5Vanderbilt Institute for Infection, Immunology, and Inflammation, Vanderbilt University, Nashville, United States

Abstract Wolbachia are the most widespread bacterial endosymbionts in animals. Within arthropods, these maternally transmitted bacteria can selfishly hijack host reproductive processes to increase the relative fitness of their transmitting females. One such form of reproductive parasitism called male killing, or the selective killing of infected males, is recapitulated to degrees by transgenic expression of the prophage WO-mediated killing (wmk) gene. Here, we characterize the genotype-phenotype landscape of wmk-induced male killing in D. melanogaster using transgenic expression. While phylogenetically distant wmk homologs induce no sex-ratio bias, closely-related homologs exhibit complex phenotypes spanning no death, male death, or death of all hosts. We demonstrate that alternative start codons, synonymous codons, and notably a single synonymous nucleotide in wmk can ablate killing. These findings reveal previously unrecognized features of transgenic wmk-induced killing and establish new hypotheses for the impacts of post-transcriptional processes in male killing variation. We conclude that synonymous sequence changes are not necessarily silent in nested endosymbiotic interactions with life-or-death consequences.

Editor’s evaluation
This study identified the genetic mechanisms underlying sex-ratio distortion through male-killing in Drosophila melanogaster flies infected with the endosymbiont Wolbachia. The endosymbiont carries the prophage WO, which is the center of interest in this study. The key result of this study is that a synonymous mutation in a prophage gene can explain the differences between sex-ratio distorting and not distorting symbionts. The finding, that a synonymous SNP plays a key role is not entirely novel in biology, but there are only few examples known of this type of genotype-phenotype association.

Introduction Wolbachia are maternally transmitted, obligate intracellular bacteria primarily residing in the cells of germline tissues in many arthropod species worldwide (Hurst and Frost, 2015; Taylor et al., 2018). To facilitate their spread through the host matriline, these bacteria hijack sex ratios, sex determination systems, and embryonic viability to cause various reproductive parasitism phenotypes (Yen and Barr, 1971; Hurst et al., 1999; Bouchon et al., 1998; Hunter, 1999). One such phenotype is male killing, whereby sons of infected females are selectively killed (Hurst et al., 1999; Hurst and Jiggins,
Male killing selfishly drives the spread of the bacteria when, for instance, brothers and sisters compete for limited resources, and male lethality yields females (the transmitters) a relative fitness benefit due to reduced competition with fewer siblings (Jaenike et al., 2003; Unkless and Jaenike, 2012; Skinner, 1985; Hurst, 1997). This form of reproductive parasitism can have profound impacts on host behavior and evolution, including a sex role reversal (Jiggins et al., 2000), host nuclear genome changes to resist the phenotype (Majerus and Majerus, 2010; Hornett et al., 2006), and potentially host population extinction (Groenenboom and Hogeweg, 2002). Population modelling also suggests male killing may be deployed as a control method to crash the population size of arthropod pests and disease vectors (Berec et al., 2016).

Many lines of evidence suggest a complex relationship between male-killing genotype and phenotype. For example, male killing can be suppressed by host background (Majerus and Majerus, 2010; Hornett et al., 2006; Jaenike, 2007; Mitsuhashi et al., 2011), where male hosts exhibit resistance to the phenotype even in the presence of Wolbachia. In addition, Wolbachia that do not cause male killing in one host species can cause male killing in a second host species upon introgression or transfer to a naive host (Jaenike, 2007; Fujii et al., 2001). Furthermore, some strains induce death at different host developmental stages (embryo vs. larvae) or across different sex determination systems (ZW lepidopterans, XY dipterans, XO arachnids), and the number of surviving males may vary widely (Riparbelli et al., 2012; Sasaki et al., 2002; Zeh et al., 2005). Such findings specify that the expression of male killing, and particularly the genotype-phenotype relationship in these symbioses, is more complex than simply the presence or absence of a male-killing gene.

We recently identified a male-killing candidate gene, wmk (WO-mediated killing), from prophage WO in the wMel Wolbachia strain of Drosophila melanogaster based on comparative genomic, transgenic, and cytological approaches (Bordenstein and Bordenstein, 2016; Perlmutter et al., 2019). It is a putative transcription factor with two predicted helix-turn-helix (HTH) DNA-binding domains, and transgenic expression in D. melanogaster embryos recapitulates many cytological and molecular aspects of male killing including accumulation of DNA damage in male embryos that overlaps with sites of dosage compensation activity (Riparbelli et al., 2012; Perlmutter et al., 2019; Harumoto et al., 2018). The wmk gene and two cytoplasmic incompatibility factor (cif) genes that underlie cytoplasmic incompatibility (a parasitism phenotype whereby offspring die in crosses between infected males and uninfected females) occur nearby specifically in the eukaryotic association module (EAM) of prophage WO (Bordenstein and Bordenstein, 2016; Perlmutter et al., 2019; LePage et al., 2017), which refers to the phage genome that is inserted into the bacterial chromosome. The EAM is common in WO phages across several Wolbachia strains (Masui et al., 2001; Wu et al., 2004) and is rich in genes that are homologous to eukaryotic genes or annotated with eukaryotic functions (Bordenstein and Bordenstein, 2016). As such, the expression of reproductive parasitism genes from the EAM and tripartite interactions between phage WO, Wolbachia, and eukaryotic hosts are central to Wolbachia’s ability to interact with and modify host reproduction.

The discovery of wmk has now enabled investigation of the impacts of genetic variation on the transgenic male-killing phenotype. Although the role of wmk in male killing has not yet been assessed in nature, evidence based on native wmk genotypes and transgenic wmk expression in flies suggest refined intricacy to the interactions between genotype and phenotype. For example, the native Wolbachia wmk homologs in three closely-related strains (wMel of D. melanogaster, wRec of D. recens, and wSuzi of D. suzukii) have nearly identical nucleotide sequences that vary by one to three SNPs (Perlmutter et al., 2019), and yet these strains induce different forms of reproductive parasitism. wMel and wRec induce cytoplasmic incompatibility, wRec also causes male killing in some strains of a sister species, and wSuzi is not known to induce any reproductive phenotype (Jaenike, 2007; Hamm et al., 2014; Hoffmann, 1988). Previous transgenic testing also demonstrated that addition of nine additional amino acids to the N-terminal region of the Wmk protein at the site of a putative alternative start codon ablates the phenotype (Perlmutter et al., 2020). Therefore, the Wmk N-terminus is particularly sensitive to genetic alterations, which is notable since small modifications such as protein tags in this location are often common additions to proteins that do not typically interfere with function. Furthermore, D. melanogaster does not harbor Wolbachia known to cause male killing, although transgenic wmk expression recapitulates the phenotype in this host (Perlmutter et al., 2019), and the wMel strain is most closely-related to the wRec male-killing strain (Metcalf et al., 2014). In addition, some D. melanogaster contain...
male-killing Spiroplasma symbionts whose plasmid gene Spaid can also transgenically recapitulate male killing (Harumoto and Lemaitre, 2018). In summary, the evidence for Wolbachia and wmk transgenic male killing suggests an intricate and multifaceted genotype-phenotype landscape.

Building on this background, several key questions emerge: Do closely-related or distantly-related homologs induce male killing? How sensitive is the wmk phenotype to small or large genetic changes? And how adapted are wmk homologs to their arthropod hosts? Here, we evaluate codon-optimized wmk homologs that span a spectrum of genetic divergence (Figure 1), including homologs (i) from distantly-related hosts, (ii) with putative upstream alternative start codons, (iii) with a single amino acid change, or (iv) with one or more synonymous codon changes, all in the D. melanogaster host. In this way, we aimed to investigate how a variety of genetic alterations to wmk, from a single synonymous nucleotide to the gene level, affect the phenotype. In particular, we hypothesized that divergent strains would not induce the phenotype due to co-adaptation with a distantly-related host. We also anticipated that the alternative start codons would inhibit function based on previous results (Perlmutter et al., 2020). In contrast, we hypothesized that neither a single amino acid change nor synonymous codon changes would alter the male-killing phenotype. We report that while distant homologs do not cause a male-killing phenotype in this host, single amino acid and synonymous nucleotide changes remarkably do alter the phenotype. Thus, wmk male killing is sensitive to the full spectrum of genetic alterations at a fine-scale level not previously recognized. Notably, synonymous sequence changes and post-transcriptional processes appear to play a role in controlling the genotype-phenotype relationships that underpin wmk male killing.

Results
Closely-related homologs of wmk
We tested homologs from strains related to wMel that occur in several groups of species in the Drosophila genus, along with two from Order Lepidoptera (Figure 1—figure supplement 1). The
phylogenetic clades of both the native Wolbachia genes and their transgenes are similar to those of the host species. Roughly divided, the transgenes group into two clusters that offer a range of genetic divergence for evaluating genotype-phenotype relationships (Figure 1—figure supplement 1) - those distantly related to the wMel wmk transgene (less than 90% codon-optimized nucleotide sequence similarity to transgenic wMel wmk) and those closely related (greater than 90% identity, from wRec, wSuzi, and an HA-tagged wMel homolog).

The closely-related wmk homologs to wMel include those from Wolbachia strains (i) wSuzi of the fruit pest species D. suzukii that has no confirmed reproductive phenotype but notably occurs in populations with female-biased sex ratios (Hamm et al., 2014; Drummond et al., 2019) and (ii) wRec of D. recens that kills males when introgressed into its sister species, D. subquinaria (Jaenike, 2007). The natural wSuzi wmk homolog has one synonymous single nucleotide polymorphism (SNP) compared to the wMel wmk reference, thus yielding the same amino acid sequence. The natural wRec wmk homolog has three SNPs, one of which is non-synonymous relative to wMel wmk, located in the first HTH DNA-binding domain (Figure 2A). These transgenes were codon optimized for expression in D. melanogaster. We hypothesized that both transgenes would induce a biased sex ratio comparable to wmk when expressed in D. melanogaster. However, transgenic expression of both homologs unexpectedly resulted in death of all expressing flies, both male and female (Figure 2B). In addition, we simultaneously tested a transgene of wRec of D. recens that kills males when introgressed into its sister species, D. subquinaria, and wSuzi and an HA-tagged wMel homolog results in the biased sex ratio previously reported (Perlmutter et al., 2019), while two closely-related transgenes from strains that at least associate with female-biased host sex ratios yielded an all-killing phenotype.

Gene expression similarities
To assess if variation in gene expression underpins the phenotypic variation, we measured the transcript levels of the transgenes in embryos 4–5 h AED (after egg deposition) when wmk kills males (Perlmutter et al., 2019). Gene expression levels are not significantly different across wSuzi, wMel, wRec, or HA tag wMel wmk transgenes, indicating that transcript levels do not account for the phenotypic differences (Figure 2C). An alternative hypothesis is that transgenic expression of the wSuzi and wRec homologs impacts native expression of a host gene in D. melanogaster that causes male and female lethality. For instance, transgenic expression of the DNA-binding dosage compensation gene, male-specific lethal 2 (msl-2), can induce total lethality with male-killing Spiroplasma (Cheng et al., 2016). Based on this, we quantified msl-2 transcript levels in embryos expressing wMel wmk (sex ratio bias), wSuzi wmk (all expressing hosts die), and wBif wmk (no killing or sex ratio bias) and found that msl-2 levels were comparable across all genotypes and phenotypes (Figure 2D).

RNA structural model variation
We next considered mRNA secondary structural differences of the transgene transcripts as a factor explaining the observed phenotypic variation. Modeling RNA structures showed they were substantially different, even in the case of few to no amino acid level changes from one sequence to another. Notably, some RNA structural features grouped by phenotype, such as the location of the start codon (Figure 2E). Two transgenic transcripts, wMel wmk and HA-wmk, cause a sex ratio bias and have start codons in the middle of the structure, while the other transgenes, wSuzi wmk and wRec wmk, that kill all flies have start codons on outer loops or hairpins. Although caution is warranted with this predicted structural analysis, mRNA secondary structure could explain some phenotypic outcomes of closely-related wmk holomologs. Native wMel wmk also has a start codon on an outer loop but does not induce an ‘all killing’ phenotype like the wSuzi and wRec wmk transgenes, which may be due to lower expression of native genes (Perlmutter et al., 2019) and/or other structural differences.

Distantly-related homologs of wmk
To determine if wmk homologs from distantly-related strains induce a biased sex ratio in D. melanogaster, we transgenically expressed four codon-optimized homologs from known male-killing strains of Wolbachia: the wBo11b strain from Hypolimnas bolina butterflies (Dyson et al., 2002), the wBif strain from Drosophila bifasciata flies (Riparbelli et al., 2012), the wCaub strain from Cadra cautella
Figure 2. Transgenic expression of closely related \textit{wmk} homologs causes male-killing and all-killing phenotypes in \textit{D. melanogaster}. (A) Schematic of \textit{wMel}, \textit{wSuzi}, and \textit{wRec \textit{wmk} native nucleotide sequences. The blue tick mark indicates a non-synonymous nucleotide difference. Black tick marks indicate synonymous nucleotide changes. Numbers indicate nucleotide position across the entire 912 nucleotide sequence. (B) Sex ratios of adult flies are

Figure 2 continued on next page
shown for expressing (Act5c-Ga4) and non-expressing (CyO) embryonic offspring. Each sample point represents the adult offspring (N = 50–157, mean 86) produced by a replicate family of ten mothers and two fathers, with expressing and non-expressing flies of a given genotype being siblings. Bars represent the mean sex ratio. Statistics are based on a Kruskal-Wallis, one-way ANOVA followed by Dunn’s correction across either expressing or non-expressing flies. wRec and wSuzi wmk have no points in the expressing category due to death of most or all males and females. HA-wmk contains a 3 X HA tag in the linker region between the two helix-turn-helix domains.

This experiment was performed twice. Data and statistical outputs are available in Figure 2—source data 1 and Figure 2—source data 2, respectively. (C) Gene expression in embryos 4–5 h AED of each indicated wmk transgene from (B), relative to Drosophila housekeeping gene, rp49. There is no significant difference in expression based on a Kruskal-Wallis one-way ANOVA followed by Dunn’s correction. Data and statistical outputs are available in Figure 2—source data 3 and Figure 2—source data 4, respectively. (D) Gene expression in embryos 4–5 h AED of the host msl-2 dosage compensation gene relative to rp49 under simultaneous expression of the indicated transgene. There is no significant difference in expression based on a Kruskal-Wallis one-way ANOVA followed by Dunn’s correction. Data and statistical outputs are available in Figure 2—source data 5 and Figure 2—source data 6, respectively. (E) Predicted RNA secondary structures of native wMel wmk and several transgene strains. Black arrows point to the location of the start codon within each structure. The online version of this article includes the following source data and figure supplement(s) for figure 2:

Source data 1. Data for sex ratios of closely-related homologs corresponding to Figure 2B.
Source data 2. Statistical output of Kruskal-Wallis test corresponding to sex ratios of closely-related homologs in Figure 2B.
Source data 3. Data for qPCR of closely-related transgenes corresponding to Figure 2C.
Source data 4. Statistical output of Kruskal-Wallis test corresponding to qPCR for transgene expression in Figure 2C.
Source data 5. Data for qPCR of msl-2 expression with transgene expression corresponding to Figure 2D.
Source data 6. Statistical output of Kruskal-Wallis test corresponding to qPCR for msl-2 expression in Figure 2D.
Figure supplement 1. wRec and wSuzi transgenes expressed with an alternative start codon lose their transgenic phenotypes.
Figure supplement 1—source data 1. Data for sex ratios of 5’ alternative start codon transgene expression corresponding to Figure 2—figure supplement 1.
Figure supplement 1—source data 2. Statistical output of Kruskal-Wallis test corresponding to sex ratios of 5’ alternative start codon transgene expression in Figure 2—figure supplement 1.

moths (Sasaki et al., 2002), and the wlnn strain from D. innubila flies (Dyer and Jaenike, 2004) (same gene sequence as the wBor male-killing strain from D. borealis flies Carson, 1956; Figure 1—figure supplement 1). While transgene expression of wMel wmk induces a biased sex ratio (~one third of expressing males die), none of the more distantly-related wBol1b, wBif, wCaub, or wlnn/wBor wmk homologs yield a biased sex ratio, demonstrating that they do not recapitulate male killing when transgenically expressed in D. melanogaster under the conditions tested (Figure 3).

Alternative start codon variation and male killing
Relevant to the studies here, we previously provided evidence that some strains contain alternative start codons upstream of the annotated start for wmk, and these upstream regions are expressed in the wMel strain (Perlmutter et al., 2020). When wMel wmk was transgenically expressed with the most likely upstream start codon, the phenotype was lost, and no biased sex ratio resulted. We also showed that some non-male-killing strains tended to have more of the alternative start codons (Perlmutter et al., 2020). To determine if wRec and wSuzi wmk transgene phenotypes are similarly sensitive to transcript changes, we expressed them with upstream codons that are native to each of their genetic sequences. As previously observed with other homologs, they lost their killing phenotype with only nine amino acids added to the 5’ end of the gene, despite being smaller than many commonly used protein tags (Figure 2—figure supplement 1). All expressing flies survived with no sex ratio bias. Returning to the RNA structure models, we find that simply adding the corresponding nucleotides at the 5’ end of each homolog resulted in several predicted differences in RNA secondary structure for each transgene compared to the structures without the additional 5’ nucleotides (Figure 2—figure supplement 1). This includes additional loops and different predicted placement of the start codon.
Finally, to identify particular nucleotides that may account for phenotypic variation among the homo-
logs, we aligned the sequences of the four transgenes in Figure 2 and investigated codons that clus-
tered by phenotype (sex ratio bias for wMelwmk and HA-wmk, or all killing for wRec and wSuzi wmk).
Across the length of the genes (and excluding the HA tag), there were only two codon differences: one
at the sixteenth amino acid position and another near the end of the gene. As previous work demon-
strated that changes at the 5’ end of this gene affect phenotype (Perlmutter et al., 2020) and since
approximately the first 10 codons in model prokaryotic genes are known to substantially affect mRNA
structure and resulting translation rate (Bentele et al., 2013), we focused on the earlier codon at site
16. This codon, which codes for Serine in all homologs, segregates among sequences by phenotype.
The HA-wmk and wMel wmk transgenes, which recapitulate male killing, have a TCG codon, while the
all-killing transgenes wSuzi and wRec wmk have TCC and AGC, respectively (Figure 4A).

To functionally test if the silent changes in this Serine codon accounted for the phenotype differ-
ences, we generated three new transgenes (Figure 4B): (i) wMel wmk control with no changes
compared to the previously tested transgene (TCG codon, labeled ‘wMel wmk (new)’), (ii) wMel wmk
with three nucleotide changes in the codon that reconstitutes the AGC present in the wRec trans-
genome line (labeled ‘wRec codon’), and (iii) wMel wmk with one nucleotide change in the codon that
Figure 4. Synonymous nucleotide changes in the 16th codon position of wmk alters resulting phenotype. (A) Sequence alignment of transgenic wmk homologs. The codon farthest on the left is the fourth codon in the sequence, and the highlighted codon is the 16th, with the farthest right representing the 23rd codon, and ellipses indicating codons continuing on either side. The red box outlines where the genotypes cluster by phenotype. The ‘HA-wmk’ and ‘wMel wmk’ genotypes share the same codon in this position, and both induce male-specific death. The ‘wRec wmk’ and ‘wSuzi wmk’ genotypes both exhibit different codons from the previous two and exhibit an all-killing phenotype. Colors correlate with amino acid identity. (B) Sequence alignment of transgenes with either the wMel wmk sequence made anew (wMel wmk new), or with the 16th codon (red box) replaced with the synonymous codons from the wRec and wSuzi wmk transgenes. The colors and symbols reflect those in (A). (C) Sex ratios of adult transgenic flies are shown for expressing (Act5c-Gal4) and non-expressing (CyO) offspring that include the original transgene wMel wmk strain used in previous figures, along with the newly created identical wMel wmk (new) transgene and the additional transgenes with the single codon swapped out for the indicated codons noted in (A). wSuzi codon and wRec codon refer to the strains that have the same sequence as the wMel wmk, but with one or three silent sites changed in the single codon at the 16th amino acid position. Each sample point represents the adult offspring (N = 50–161, mean 73) produced by a replicate family of 10 mothers and two fathers, with expressing and non-expressing flies of a given genotype being siblings. Bars represent the mean sex ratio. Statistics are based on a Kruskal-Wallis one-way ANOVA followed by Dunn’s correction across either expressing or non-expressing flies. This experiment was performed twice. Data and statistical outputs are available in Figure 4—source data 1 and Figure 4—source data 2, respectively. (D) Gene expression in embryos 4–5 h AED denotes expression of each transgene relative to that of rp49. There is no significant difference in expression based on a Kruskal-Wallis one-way ANOVA followed by Dunn’s correction across either expressing or non-expressing flies. Gene expression data and statistical outputs are available in Figure 4—source data 3 and Figure 4—source data 4, respectively. (E) Predicted RNA secondary structures are shown for the wMel wmk transcript compared to both of the wRec or wSuzi codon transgenes exhibiting slight structural differences. The structure for transgene wMel wmk is included again from Figure 3 for ease of Figure 4 continued on next page
Figure 4 continued

comparison. Black arrows point to the location of the start codon within each structure. Black circles highlight a key area of difference between the structures, with a stem absent in the wSuzi codon strain, and different base pair match probabilities calculated for each as indicated by color (blue to red, low to high probability). Within the black circle, the wSuzi codon transgene structure is missing a predicted stem that the others have. The stem in the wRec codon line, while present, has a weaker prediction as noted by the cooler colors, so there may be structural differences compared to the wMel wmk model.

The online version of this article includes the following figure supplement(s) for figure 4:

Source data 1. Data for sex ratios from expression of transgenes with single codon changes corresponding to Figure 4C.
Source data 2. Data for qPCR from expression of transgenes with single codon changes corresponding to Figure 4D.
Source data 3. Data for qPCR from expression of transgenes with single codon changes corresponding to Figure 4D.
Source data 4. Statistical output of Kruskal-Wallis test corresponding to qPCR from expression of transgenes with single codon changes in Figure 4D.

reconstitutes the TCC present in the wSuzi transgene line (labeled ‘wSuzi codon’). When these three otherwise identical genes were transgenically expressed, the wMel wmk (new) transgene with no changes caused a biased sex ratio as expected; however, and remarkably, expression of transgenes with two different Serine codons ablated the phenotype and resulted in non-biased sex ratio with normal numbers of expressing flies (Figure 4C). This ablation occurs even though transcript levels remain similar across all transgenes and despite sequencing confirmation of the single codon differences (Figure 4D). Thus, a minimum of one single synonymous site change in the 5’ region was sufficient to alter the sex ratio phenotype. However, while the nucleotide changes in the codon changed the phenotype, they did not recapitulate the all-killing phenotypes of their corresponding homologs. The predicted RNA secondary structures from the transgenes with the single codon changes are similar to the original wMel wmk transgene, but they differ in some aspects such as presence or absence of stems and loops and the probability score of the base pair match as indicated by color (scale of red to blue, warmer colors indicate high probability, cooler colors indicate low probability). (Figure 4E).

Discussion

Linking mutations to function is crucial in resolving the evolutionary dynamics of adaptations, especially when the functions emerge in a nested system of tripartite phage-bacteria-animal interactions. In a simple case, the direct impact of genetic divergence is functional divergence with increasing numbers of mutations leading to increased functional divergence. Here, we investigated evolutionary and molecular hypotheses related to the genotype-phenotype relationships between the prophage WO-mediated killing (wmk) gene (Perlmutter et al., 2019; Perlmutter et al., 2020) and male killing in D. melanogaster. Sequences among wmk homologs in nature can differ across divergent hosts, leading to the hypothesis that highly divergent alleles are functionally fine-tuned. Using a variety of sequences, we report three key results: (i) mutations even at the synonymous codon and single nucleotide levels can alter the male-killing phenotype, (ii) phenotype, genotype, and RNA structural variation exhibit some correlation with each other, and (iii) distantly-related homologs do not induce a male-killing phenotype in transgenic D. melanogaster, while some closely-related homologs from Drosophila species/populations with female-biased sex ratios induce all-killing phenotypes. We discuss how these findings expand and support mechanistic models of wmk male killing, emphasize important implications for transgenic assays in endosymbiont studies, and relate these findings to aspects of male killing such as phenotype switching and host resistance.

The most surprising and unanticipated result was that transgene expression of highly similar homologs and even single synonymous site changes alter phenotype, and they make the difference between life and death for some males. We tested wmk homologs with high native sequence identity compared to native wMel wmk: wRec wmk from the mushroom-feeding D. recens (native sequence has two synonymous and one non-synonymous nucleotide differences) and wSuzi wmk from the fruit crop pest D. suzukii (native sequence has one synonymous nucleotide difference). Although we anticipated similar results to wMel wmk expression, we found that transgenic expression of these genes killed all flies (Figure 2), even though the wSuzi wmk transgene produces an identical protein to transgenic wMel Wmk. Of particular note, transgenic wSuzi wmk (all killing) and the original transgenic
wMel wmk sequence (male killing) produce proteins of the exact same amino acid sequence, but they shared only a 91 % sequence identity in the codon-optimized transgene, largely due to an updated codon optimization algorithm. Codon optimization is the norm in transgenic symbiosis research under the common assumption that synonymous codons are functionally redundant. However, codon optimization algorithms often choose a codon based on factors including codon adaptation, mRNA folding, regulatory motifs, nucleotide bias, or codon correlations and biases (Plotkin and Kudla, 2011). With even a few different codons input into the algorithm as well as algorithm updates over the years, the tested transgenes had different nucleotide sequences.

Based on the aforementioned results, we next sought to assess the sensitivity of male killing to wmk transcriptional and post-transcriptional changes and asked if one synonymous codon or site change was sufficient to alter phenotype. We created three transgenic lines, each with different DNA sequences coding for the same Serine at position 16. With only these minor changes that encode an identical protein sequence to wMel Wmk, the transgenic male-killing phenotype was ablated (Figure 4). Thus, one single codon and even a single nucleotide, remarkably determined male viability. Importantly, this key result implies at the molecular level that endosymbiont phenotypes of reproductive parasites may not simply be governed by DNA sequence alone.

A key question related to our findings is how synonymous changes cause vastly different phenotypes. It is often assumed in endosymbiotic transgenic experiments that synonymous changes do not affect function, and codons for the same amino acid are functionally redundant. However, several decades of research have uncovered mounting evidence that the functional redundancy hypothesis is not always accurate (Plotkin and Kudla, 2011). Indeed, codon bias varies across species (hence the need for codon optimization) (Sharp et al., 1997; Andersson and Kurland, 1990), and there are several mechanisms through which single or few codons can influence transcription or translation. When there is a rare codon, there are fewer tRNAs available, and the translation rate may be correspondingly lowered (Li et al., 2012). Notably, there is a bias for rare codons in N-terminal regions of bacterial genes, which is likely due to their influence on mRNA structure (Gu et al., 2010). Lower 5’ RNA stability or weaker structure can also be correlated with ease of initiation and faster translation (Gu et al., 2010), and stem loop structures may affect translation rate in varying ways. Indeed, strong 5’ RNA secondary structure may inhibit ribosomal initiation or translational efficiency (de Smit and van Duin, 1990; Bettany et al., 1989), and there can be selection against SNPs that alter RNA secondary structure (Salari et al., 2013). There is also evidence for selection across the domains of life on codon usage in the first 30–60 nucleotides of a gene, likely due to their impact on mRNA structure near the site of initiation (Gu et al., 2010). Additionally, early codons with low frequencies can be beneficial by slowing elongation, thus preventing ribosome collisions, or potentially helping to recruit chaperones to the emerging peptide for proper folding (Tuller et al., 2010; Fredrick and Ibba, 2010). In addition, some specific codons across the sequence are favored for their lower likelihood of mistranslation (Qin et al., 2004). Indeed, synonymous codon differences in genes can result, for instance, in altered gene expression (Kudla et al., 2009), lower organism fitness (Agashe et al., 2013), or disease (Sauna and Kimchi-Sarfaty, 2011).

Beyond codon sequence, gene expression levels across representatives of all phenotypes are similar (Figure 2), supporting post-transcriptional differences as the source of functional variation and refuting the functional redundancy hypothesis in this context. Examining the three synonymous transgenes more closely, we find that codon usage indeed correlates with phenotype. According to the codon usage table for D. melanogaster on the Genscript website, the original wMel wmk codon at this position (TCG) has a frequency 16.7 per 1000 codons. In contrast, the wSuZi codon (TCC) has a frequency of 19.5/1000, and the wRec codon (AGC) has a frequency of 20.5/1000 (GenScript, 2020). Therefore, there is a pattern where the more common codons are present in the all-killing transgenes, and at similar frequencies, while the strain expressing partial male killing has a codon at a lower frequency. Thus, some of the phenomena described in studies on synonymous codons and function likely underlie the phenotypic variation observed here.

Given that transgenic wmk expression with silent site changes results in different phenotypes, we used software that estimated RNA secondary structures for the tested transgenes. Indeed, we found that the predicted structures correlate somewhat with phenotype, so synonymous codons may change transcript structure or other post-transcriptional features of the transcript or protein (Figure 2). Other small changes including adding a 9-codon N-terminal sequence ablated the phenotype and reversed
it to no death, despite the additional sequence being smaller than many common gene tags that typically do not interfere with function. This small change also alters predicted RNA structure (Figure 2—figure supplement 1). We were unable to assess protein levels to determine translational differences due to lack of an antibody. However, the data demonstrates that the wmk transgene is functionally sensitive to some post-transcriptional changes, at least in the N-terminal region.

What, then, do these results illuminate about transgenic wmk function and the utility of transgenic research in light of the potential for marked influences of synonymous changes on phenotype? First, as discussed, the wmk transgenic phenotypes are likely sensitive to post-transcriptional processes. This has important implications for understanding wmk in a natural male-killing context as soon as feasible techniques are available, since heterologous expression and its reliance on codon optimization may obscure our understanding of the gene’s biology. Second, since three different transgenic phenotypes (sex ratio bias, all killing, no killing) have been found so far with only a few sequences analyzed, it is possible that further testing of new codons may increase the partial male killing to a full male-killing phenotype. Therefore, it would be fruitful and may be possible to continue to uncover a transgenic sequence that fully recapitulates the phenotype to refine this system as a study tool for Wolbachia male killing. Third and critically, although the results show that there is some relationship between synonymous codons and phenotype, several points remain for further testing. For example, we cannot conclude that the particular codon tested here is responsible for phenotype alterations in other host genetic backgrounds or species. It is possible that this codon plays a functional role only in a singular host genetic context. Here, we changed wmk sequences while holding the host genetic background fixed, but the reverse is required to conclude whether or not the particular codon plays a general role in other genotypes or natural contexts. Second, due to possible coevolution, various codons may or may not yield similar functional effects across different host backgrounds, and additional synonymous sites may contribute to the male-killing phenotype. Thus, the results here illuminate a previously unrecognized need for future research on the functional impacts of synonymous substitutions in endosymbionts. Future work may focus on determining if there is one specific synonymous codon that affects the male-killing function in all cases, if a more general feature exists where alteration of any or a subset of N-terminal or other wmk codons affects function, or if the effect of synonymous changes is specific to this background.

In addition, these findings are informative with regards to the more general study of phenotypes induced by endosymbiont transgenes. It is standard practice to codon optimize genes for maximizing host expression when testing endosymbiont gene function (Perlmutter et al., 2019; LePage et al., 2017; Harumoto and Lemaître, 2018), with the assumption that synonymous codons will not alter function. However, if codon optimization potentially changes the interpretation of transgenic findings, then phenotypes should be corroborated in natural contexts once tools such as genetic editing are available in the relevant organism. Specifically, wmk should be knocked out in native contexts once it is more technically feasible. In addition, codon optimization algorithms are updated with new information periodically with the assumption that they yield improved results, although it is unclear in practice whether an algorithm is better optimized to produce results that reflect the true biology of a transgene. Future work is necessary to explore these concepts further. For example, comparisons of alleles may need to be performed with alleles identical in sequence except for any engineered differences, and the algorithm should remain constant across all transgenes that are compared to each other. Further, careful analysis and comparison of transgenic phenotypes produced by different algorithms may be necessary in some cases where the phenotype is known or expected. This approach could ensure the algorithm produces a transgenic phenotype that most closely resembles the natural phenotype. In addition, the reliability of certain molecular evolutionary signatures, such as dN/dS, may be compromised since synonymous mutations are assumed to be neutral in these analyses. These principles are particularly important for research on endosymbionts that increasingly relies on heterologous gene expression for functional studies. The smaller genomes of endosymbionts tend to have low GC content overall (McCutcheon et al., 2009), and GC content is the strongest driver of codon usage bias (Knight et al., 2001; Chen et al., 2004) and contributes to the strength of mRNA secondary structure (Plotkin and Kudla, 2011), so careful attention to the effects of synonymous changes may be of acute interest to the endosymbiont research community.
We also find interesting support for the hypothesis that there is co-adaptation between wmk homologs and their hosts since male-killing genes may be evolutionary matched for the host sex determination and molecular machinery that they manipulate. We first analyzed wmk homologs that are more distantly related to wMel wmk, including from male-killing strains wBo1b, wBif, wCaub, and wmn/wBor (Sasaki et al., 2002; Dyson et al., 2002; Dyer and Jaenike, 2004; Hurst et al., 2000; Sheeley and McAllister, 2009). They did not recapitulate male killing when transgenically expressed in D. melanogaster. The lack of a sex ratio bias is expected if either wmk does not underpin male killing in these systems or divergent homologs are required to be closely co-adapted to their hosts. This latter hypothesis is based on observations that resistance to male killing is common (evidence of a potential host-microbe arms race) (Charlat et al., 2005; Majerus and Majerus, 2010; Jaenike, 2007; Mitsuhashi et al., 2011; Hornett et al., 2008), and some strains can cause male killing anew when transferred to another, usually closely-related host (Jaenike, 2007; Fujii et al., 2001; Sasaki et al., 2002) while not causing male killing in all recipient host species (Veneti et al., 2012; Hughes and Rasgon, 2014). It is possible, then, that the wmk homologs tested induce a sex ratio bias in their natural hosts, but not in D. melanogaster.

If we make the assumption that wmk is involved in male killing in nature, which requires confirmation beyond transgenic recapitulation of the phenotype, then the results here give the basis for additional hypotheses that require further testing. First, the killing of both males and females by the closely-related wRec and wSuzi wmk homologs could indicate that the target is something shared in both males and females, but functions differently within each sex (assuming no off-target effects). Indeed, previous studies on wmk and Wolbachia male killing demonstrate a positive correlation with between dosage compensation complex (DCC) activity and DNA defects in male embryos (Riparbelli et al., 2012; Perlmutter et al., 2019; Harumoto et al., 2018). Notably, four of the five protein components of the DCC are expressed in both males and females, and it is only msl-2 that is male-specific and catalyzes formation of the complex (Lucchesi and Kuroda, 2015). Thus, there are several non-sex-specific wmk target candidates that the gene product may interact with to cause lethality in males and females. Second, protein divergence and resultant conformational changes may impact the specificity between host target and the Wmk toxin, and could underlie development of host resistance. Male-killing genes are expected to evolve rapidly within hosts in order to counteradapt host resistance mechanisms and keep evolutionary pace with the rapid evolution of sex-related genes that may be the target of a male-killing toxin (Marín and Baker, 1998; Rodríguez et al., 2007). As such, major or minor divergence in protein or transcript sequence of either the host target or the microbial toxin may underpin changes that lead to common host resistance, wmk-host coadaptation, and functional transgene differences in the foreign D. melanogaster host. Third, these findings open up the possibility of a variety of functionally relevant wmk protein or transcript conformations in nature, which could contribute to the marked diversity of Wolbachia male killing in terms of host species and sex determination systems (Hurst et al., 1999; Charlat et al., 2005; Sasaki et al., 2002; Zeh et al., 2005; Dyer and Jaenike, 2004; Hurst et al., 2000; Fialho and Stevens, 2000; Van Borm et al., 2008).

Taken together, this work reports previously unrecognized relationships for wmk-induced killing and establishes new hypotheses for the impacts of RNA structure and post-transcriptional processes in wmk-induced male killing. It also highlights several critical features for the research community regarding assumptions about the broad use of transgenes and the role of synonymous mutations in gene function. If wmk is involved in natural male killing, then this work could indicate how silent sequence changes may relate to known male-killing phenomena such as frequent host resistance or male-killing function in diverse hosts.

### Materials and methods

#### Key resources table

| Reagent type (species) or resource | Designation | Source or reference | Identifiers | Additional information |
|-----------------------------------|-------------|---------------------|-------------|------------------------|
| Gene (Wolbachia pipiens)          | WD0626      | NCBI                | NCBI:WD_RS02815 | Also known as wmk (WO-mediated killing) |

Continued on next page
| Reagent type (species) or resource | Designation | Source or reference | Identifiers | Additional information |
|-----------------------------------|-------------|--------------------|-------------|------------------------|
| Genetic reagent (D. melanogaster) | Act5c-Gal4/CyO | Bloomington Drosophila Stock Center | BDSC:3953; FlyBase FBtr0012290 | P(AyGAL4)25 |
| Genetic reagent (D. melanogaster) | WT (y1w67c23; P[CaryP]P2) | Bloomington Drosophila Stock Center | BDSC:8622; FlyBase FBtr0040535 | WT strain used in this study; P(CaryP)attP2 |
| Genetic reagent (D. melanogaster) | wMel wmk | This paper; Perlmutter et al., 2019; PMID:31504075 | | Expresses codon-optimized transgene; UAS promoter |
| Genetic reagent (D. melanogaster) | wBol1b wmk | This paper | | Expresses codon-optimized transgene; UAS promoter |
| Genetic reagent (D. melanogaster) | wBif wmk | This paper | | Expresses codon-optimized transgene; UAS promoter |
| Genetic reagent (D. melanogaster) | wCaub wmk | This paper | | Expresses codon-optimized transgene; UAS promoter |
| Genetic reagent (D. melanogaster) | wInn/wBor wmk | This paper | | Expresses codon-optimized transgene; UAS promoter; wInn and wBor wmk have same exact sequence |
| Genetic reagent (D. melanogaster) | wSuzi wmk | This paper | | Expresses codon-optimized transgene; UAS promoter |
| Genetic reagent (D. melanogaster) | wRec wmk | This paper | | Expresses codon-optimized transgene; UAS promoter |
| Genetic reagent (D. melanogaster) | HA-wmk | This paper | | Expresses codon-optimized transgene; UAS promoter; 3 X HA tag epitope in linker between HTH domains of wMel wmk |
| Genetic reagent (D. melanogaster) | 5' wRec wmk | This paper | | Expresses codon-optimized transgene; UAS promoter; Sequence has additional nine amino acids starting at upstream alternative start codon |
| Genetic reagent (D. melanogaster) | 5' wSuzi wmk | This paper | | Expresses codon-optimized transgene; UAS promoter; Sequence has additional nine amino acids starting at upstream alternative start codon |
| Genetic reagent (D. melanogaster) | wMel wmk (new) | This paper | | Expresses codon-optimized transgene; UAS promoter; Same exact sequence as wMel wmk, newly transformed strain |
| Genetic reagent (D. melanogaster) | wSuzi codon | This paper | | Expresses codon-optimized transgene; UAS promoter; Same as wMel wmk, but with 16th amino acid position using TCC Serine codon from wSuzi wmk strain |

Continued on next page
Drosophila strains and maintenance

*D. melanogaster* strains used in this study include Act5c-Gal4/CyO (BDSC 3953, ubiquitously-expressing zygotic driver), the WT background line of genotype y w^{67c23}; P[CaryP]P2 (BDSC 8622), the WD0626 (wmk) and WD0034 (control gene) transgene constructs previously described (Perlmutter et al., 2019), and several new transgene constructs. Briefly, each Wolbachia gene of interest was codon-optimized for optimal *D. melanogaster* expression using algorithms developed by GenScript Biotech (Piscataway, NJ). These sequences were then synthesized as DNA nucleotides and cloned by GenScript into the pTIGER plasmid (Ferguson et al., 2012). The pTIGER vector is a UASp-based plasmid optimized for germline expression and uses PhiC31 integrase (Groth et al., 2004) for targeted integration into the *D. melanogaster* genome. It also includes UAS promoters for inducible expression with a Gal4 driver (Southall et al., 2008) as well as a w~+~ red eye marker for transformant screening. The plasmids were then sent to BestGene (Chino Hills, CA), which performed injections of the vectors into embryos of the BDSC 8622 background line with PhiC31 integrase to integrate the vector into the attP2 insertion site in the genome. Successful transformants were selected based on

| Reagent type (species or resource) | Designation | Source or reference | Identifiers | Additional information |
|-----------------------------------|-------------|---------------------|-------------|------------------------|
| Genetic reagent (*D. melanogaster*) | wRec codon | This paper | | Expresses codon-optimized transgene; UAS promoter; Same as wMel wmk, but with 16th amino acid position using AGC Serine codon from wRec wmk strain |
| Recombinant DNA reagent | pTIGER (plasmid) | *Ferguson et al., 2012,* PMID:22328499 | | Modified pUASP plasmid for enhanced germline expression under Gal4/UAS control |
| Sequence-based reagent | Rp49_F | This paper | PCR primers | CGGTACGGATCGAACAAGC |
| Sequence-based reagent | Rp49_R | This paper | PCR primers | CTTCGGCTCTTCGAGGAGA |
| Sequence-based reagent | wmk_homologs_opt_F | This paper | PCR primers | CTGTAAGCCATTGCAGAGCC |
| Sequence-based reagent | wmk_homologs_opt_R | This paper | PCR primers | TCACCATGCTCTGAGGAGA |
| Sequence-based reagent | Msl-2_F | This paper | PCR primers | GAATTAAGCGGGTCTAAGATGTA |
| Sequence-based reagent | Msl-2_R | This paper | PCR primers | GTATGCCTGTCGGCATGAT |
| Commercial assay or kit | Direct-zol RNA MiniPrep Kit | Zymo | | R2051 |
| Commercial assay or kit | Superscript VILO cDNA Synthesis Kit | ThermoFisher | | 11754050 |
| Chemical compound, drug | DNase, RNase-free | Ambion, Life Technologies | | AM2222 |
| Chemical compound, drug | iTaq Universal SYBR Green Mix | Bio-Rad | | 1725120 |
| Software, algorithm | GraphPad Prism 8 | GraphPad Prism 8 | | RRID:SCR_002798 |
| Software, algorithm | Geneious Pro v.2019.2; Geneious Pro v.2020.2.4 | Geneious | | RRID:SCR_010519 |
| Software, algorithm | jModelTest | jModelTest | | RRID:SCR_015244 |
| Software, algorithm | RNAfold WebServer | University of Vienna, *Gruber et al., 2008,* Lorenz et al., 2011 | PMID:18424795; PMID:2215189 | http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi |

**Reagents source from this paper may be obtained from Bordenstein lab.**
the red eye marker, and each transgene line is descended from the offspring of a single transformant (isofemale).

D. melanogaster were reared on standard cornmeal, molasses, and yeast (CMY) media (5% w/v cornmeal (Quaker, Chicago IL), 1.875% w/v yeast (Red Star Yeast, Milwaukee WI), 7.8% v/v molasses (Sweet Harvest Foods, Rosemount MN), 0.5% w/v type II Drosophila agar (Genesee Scientific, San Diego CA), 0.056% w/v tegosept (Genesee Scientific, San Diego CA), and 0.39% v/v propionic acid (Sigma Aldrich)). Stocks were maintained at 25 °C with virgin flies stored at room temperature. During virgin collections, stocks were kept at 18 °C overnight and 25 °C during the day. All flies were kept on a 12 hr light/dark cycle.

**Sex ratio assays**

To assess the effect of transgene expression on adult sex ratios (measurement of male killing), sex ratio assays were performed as previously described (Perlmutter et al., 2019). Briefly, twenty biological replicates of 10 uninfected, 4- to 7-day-old virgin, female Act5c-Gal4/CyO driver flies and two uninfected, 1- to 2-day-old virgin, male transgene flies were each set up in vials with CMY media. Individuals were randomly allocated to each vial after all females or males of a given genotype were mixed together. They were left on the media to lay eggs for 4 days at 25 °C with a 12 hr light/dark cycle, at which point adults were discarded. The vials are then left at 25 °C until the offspring are counted.

After 9 days of adult offspring emergence, they were scored for both sex and expression (red eye color from Act5c-Gal4 chromosome) or non-expression (curly wings from CyO balancer chromosome). The number of adult offspring per vial across all experiments ranges from 50 to 170, with a mean of 120 (ranges and means per experiment are included in figure captions). Any vials with fewer than 50 adult offspring were removed from the analysis, as this indicates either abnormally poor egg laying or hatching (typically 0–2 vials per group). In addition, vials with no adult emergence, while others of the same genotype had typical levels of offspring, were also excluded (typically 2–3 vials per group). Results were graphed in GraphPad Prism 8.4.0, which applies a ‘Standard Plot Appearance’ correction for visibility of data distribution where the width of distribution of points is proportional to the number of points at that y-value.

**RNA secondary structures**

RNA secondary structures were generated by uploading the nucleotide sequences of the indicated gene to the RNA fold web server (Gruber et al., 2008; Lorenz et al., 2011). The structures shown are the graphical outputs of the MFE (minimum free energy) secondary structures. Colors indicate base pair probabilities, from blue to red, with blue indicating a probability of 0 and red indicating a probability of 1.

**Gene expression**

Gene expression was measured in Drosophila embryos aged 4–5 hr AED. Each point represents a biological replicate with the RNA of 30 pooled embryos from crosses between a unique set of 60 uninfected, 4- to 7-day-old virgin, female Act5c-Gal4/CyO driver flies and 12 uninfected, 1- to 2-day-old virgin, male transgene flies of the indicated genotype. Each point represents a biological replicate from different bottles. Individuals were randomly allocated to each bottle after all females or males of a given genotype were mixed together. Each collection chamber consisted of a grape juice agar plate with yeast in an eight oz round bottom bottle, and flies. These were placed in a 25 °C incubator overnight (16 hr). Then, the plates were swapped with fresh ones. The flies were allowed to lay eggs for 1 hr. The plates were then left at 25 °C for an additional 4 hr to age them to be 4–5 hr old (the estimated time of male death in wmk crosses). Embryos were then gathered in groups of 30 (each group from a unique bottle/biological replicate) and flash frozen in liquid nitrogen. RNA was extracted using the Direct-zol RNA MiniPrep Kit (Zymo), RNase-free DNase (Ambion, Life Technologies), cDNA was generated with SuperScript VILO (Invitrogen), and RT-qPCR was run using iTaq Universal SYBR Green Mix (Bio-Rad). qPCR was performed on a Bio-Rad CFX-96 Real-Time System. Primers are listed in the Key Resources Table. Conditions were as follows: 50 °C 10 min, 95 °C 5 min, 40 x (95 °C 10 s, 55 °C 30 s), 95 °C 30 s. Differences in gene expression were done by calculating 2-ΔΔct (difference in ct values of two genes of interest). Data points were excluded if a sample had low-quality cDNA that did not
amplify in qPCR. Data points of each biological replicate are measured as the mean of two technical replicates from each sample.

**Phylogenetic trees**

The nucleotide phylogenetic trees of host COI genes and wmk native gene or transgene sequences were inferred based on a MUSCLE alignment in Geneious Prime 2020.2.4 followed by stripping all sites with gaps. The resulting 652, 690, 686 nucleotide base pair alignments (respectively) were then analyzed via jModelTest 2.1.10 v20160303 (Darriba et al., 2012Guindon and Gascuel, 2003). The AICc-corrected best model, JC, was predicted for all three alignments and was used to build the trees using the MrBayes (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) Geneious plugin with the JC69 model (Jukes and Cantor, 1969) and equal rate variation.

**Transgene sequence alignments**

The sequence alignments of different wmk transgenes in Figure 2 and Figure 4 were conducted in Geneious Pro v.2019.2 using a MUSCLE alignment. Black bars in Figure 2 indicate sequence mismatches compared to the wMel wmk transgene reference sequence with any gaps stripped. Codons in Figure 4 are colored by amino acid.

**Statistical analyses**

Sample sizes for experiments were based on previous publications, which demonstrated repeatability in relative differences between treatment groups. Each experiment was completed twice, and statistical tests were applied to both to confirm repetition in differences between treatment groups. In all plots, the first experiment is the representative one shown. For sex ratios, we tested different sample sizes for reliability in sex ratio measurements, and found that 20 biological samples per group, with 10 females and 2 males per sample, resulted in consistently replicable data, which is the standard we apply here. qPCR data was approached similarly, by previous work in the lab demonstrating that with embryos 4–5 h AED, we are able to get consistent, high-quality, replicable data with 30 embryos per sample, and at least eight samples per group.

All statistical analyses for sex ratios and were performed using GraphPad Prism eight software. For sex ratios, a non-parametric Kruskal-Wallis one-way ANOVA followed by Dunn’s multiple comparisons test was applied to all gene-expressing categories, followed by the same test but on all non-expressing categories. For gene expression, groups were compared using the same Kruskal-Wallis one-way ANOVA with Dunn’s multiple comparisons test. Significant results are indicated with * symbols in figures, with accompanying $p$ values in captions. Any comparisons with no symbol are nonsignificant. Full statistical information and outputs for all sex ratio and qPCR data is available in the Source Data file.

**Acknowledgements**

This work was supported by National Institutes of Health (NIH) grant R21 AI133522 to SRB, the Vanderbilt Microbiome Innovation Center, and NIH grants F31 AI143152, K-INBRE P20 GM103418, and National Science Foundation (NSF) Postdoctoral Research Fellowship (PRFB) DBI 2109772 to JIP. We would also like to thank RL Unckless for helpful comments on phylogenies.

**Additional information**

**Competing interests**

Jessamyn I Perlmutter, Seth R Bordenstein: Is listed as an author on a patent related to the use of wmk in vector control. US Patent 20210000092 16/982708. The other author declares that no competing interests exist.
Funding

| Funder                               | Grant reference number | Author                  |
|--------------------------------------|------------------------|-------------------------|
| National Institutes of Health        | R21 AI133522           | Seth R Bordenstein      |
| National Institutes of Health        | F31 AI143152           | Jessamyn I Perlmutter   |
| Vanderbilt Microbiome Innovation Center | General Funds        | Seth R Bordenstein      |
| National Institutes of Health        | P20 GM103418           | Jessamyn I Perlmutter   |
| National Science Foundation          | DBI 2109772            | Jessamyn I Perlmutter   |

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Jessamyn I Perlmutter, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review and editing; Jane E Meyers, Investigation, Writing – original draft; Seth R Bordenstein, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing

Author ORCIDs

Jessamyn I Perlmutter [http://orcid.org/0000-0002-9789-4674](http://orcid.org/0000-0002-9789-4674)
Seth R Bordenstein [http://orcid.org/0000-0001-7346-0954](http://orcid.org/0000-0001-7346-0954)

Decision letter and Author response

Decision letter [https://doi.org/10.7554/eLife.67686.sa1](https://doi.org/10.7554/eLife.67686.sa1)
Author response [https://doi.org/10.7554/eLife.67686.sa2](https://doi.org/10.7554/eLife.67686.sa2)

Additional files

Supplementary files
- Transparent reporting form
- Source data 1. Source data for all graphical data sets and statistical tests performed for this study.

Data availability

All data generated or analyzed during this study are included in the manuscript and supporting files. Source data files have been provided for Figures 3-6.

References

Agashe D, Martinez-Gomez NC, Drummond DA, Marx CJ. 2013. Good codons, bad transcript: large reductions in gene expression and fitness arising from synonymous mutations in a key enzyme. *Molecular Biology and Evolution* 30: 549–560. DOI: [https://doi.org/10.1093/molbev/msss273](https://doi.org/10.1093/molbev/msss273), PMID: 23223712

Andersson SG, Kurland CG. 1990. Codon preferences in free-living microorganisms. *Microbiological Reviews* 54: 198–210. DOI: [https://doi.org/10.1128/mr.54.2.198-210.1990](https://doi.org/10.1128/mr.54.2.198-210.1990), PMID: 2194095

Bentele K, Saffert P, Rauscher R, Ignatova Z, Blüthgen N. 2013. Efficient translation initiation dictates codon usage at gene start. *Molecular Systems Biology* 9: 675. DOI: [https://doi.org/10.1038/msb.2013.32](https://doi.org/10.1038/msb.2013.32), PMID: 23774758

Berec L, Maxin D, Bernhauerová V, Rohr J. 2016. Male-killing bacteria as agents of insect pest control. *Journal of Applied Ecology* 53: 1270–1279. DOI: [https://doi.org/10.1111/1365-2664.12638](https://doi.org/10.1111/1365-2664.12638)

Bettany AJ, Moore PA, Cafferkey R, Bell LD, Goody AR, Carter BL, Brown AJ. 1989. 5′-secondary structure formation, in contrast to a short string of non-preferred codons, inhibits the translation of the pyruvate kinase mRNA in yeast. *Yeast* 5: 187–198. DOI: [https://doi.org/10.1002/yea.320050308](https://doi.org/10.1002/yea.320050308), PMID: 2660464

Bordenstein SR, Bordenstein SR. 2016. Eukaryotic association module in phage WO genomes from Wolbachia. *Nature Communications* 7: 13155. DOI: [https://doi.org/10.1038/ncomms13155](https://doi.org/10.1038/ncomms13155), PMID: 27712237

Perlmutter et al. eLife 2021;0:e67686. DOI: [https://doi.org/10.7554/eLife.67686](https://doi.org/10.7554/eLife.67686)
Bouchon D, Rigaud T, Juchault P. 1998. Evidence for widespread Wolbachia infection in isopod crustaceans: molecular identification and host feminization. *Proceedings. Biological Sciences* **265**: 1081–1090. DOI: https://doi.org/10.1098/rspb.1998.0402, PMID: 9684374

Carson H. 1956. A female producing strain of *D. borealis* Patterson. *Drosoph Inf Serv* **30**: 109–110.

Charlat S, Hornett EA, Dyson EA, Ho PPy, Loc NT, Schiltuizen M, Davies N, Roderick GK, Hurst GDD. 2005. Prevalence and penetrance variation of male-killing Wolbachia across Indo-Pacific populations of the butterfly *Hypolimnas bolina*. *Molecular Ecology* **14**: 3525–3530. DOI: https://doi.org/10.1111/j.1365-294X.2005.02678.x, PMID: 16156820

Chen SL, Lee W, Hottes AK, Shapiro L, McAdams HH. 2004. Codon usage between genomes is constrained by genome-wide mutational processes. *PNAS* **101**: 3480–3485. DOI: https://doi.org/10.1073/pnas.0307827100, PMID: 14990797

Cheng B, Kuppanda N, Aldrich JC, Akbari OS, Ferree PM. 2016. Male-killing Spiroplasma alters behavior of the male *Drosophila*. *Lee W, Hottes AK, Shapiro L, McAdams HH. 2004. Codon usage between genomes is constrained by genome-wide mutational processes. PNAS* **101**: 3480–3485. DOI: https://doi.org/10.1073/pnas.0307827100, PMID: 14990797

Groenenboom MAC, Hogeweg P. 2002. Space and the persistence of male-killing endosymbionts in insect populations. *Proceedings. Biological Sciences* **269**: 2509–2518. DOI: https://doi.org/10.1098/rspb.2002.2197, PMID: 12573064

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772. DOI: https://doi.org/10.1038/nmeth.2109, PMID: 22847109

de Smit MH, van Duin J. 1990. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *PNAS* **87**: 7668–7672. DOI: https://doi.org/10.1073/pnas.87.19.7668, PMID: 2217199

Drummond F, Ballman E, Collins J. 2019. Population Dynamics of Spotted Wing *Drosophila* (*Drosophila suzukii* (Matsumura)) in Maine Wild Blueberry (*Vaccinium angustifolium* Aiton). *Trends in Genetics* **34**: 475–479. DOI: https://doi.org/10.1016/j.tig.2019.06.011

Gu W, Zhou T, Wilke CO, Snel B. 2010. A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLOS Computational Biology* **6**: e1000664. DOI: https://doi.org/10.1371/journal.pcbi.1000664

Gundersen, M. 2008. How the sequence of a gene can tune its translation. *Cell* **141**: 227–229. DOI: https://doi.org/10.1016/j.cell.2010.03.033, PMID: 20403320

Fujii Y, Kageyama D, Hoshizaki S, Ishikawa H, Sasaki T. 2001. Transfection of Wolbachia in Lepidoptera: the site-specific recombination site in prokaryotes and eukaryotes. *PLOS Computational Biology* **6**: e1000664. DOI: https://doi.org/10.1371/journal.pcbi.1000664

Fialho RF, Stevens L. 2000. Male-killing *Wolbachia* in a flour beetle. *Proceedings of the Royal Society of London. Series B* **267**: 1469–1473. DOI: https://doi.org/10.1098/rspb.2000.1166

Fujii Y, Kageyama D, Hoshizaki S, Ishikawa H, Sasaki T. 2001. Transfection of Wolbachia in Lepidoptera: the site-specific recombination site in prokaryotes and eukaryotes. *PLOS Computational Biology* **6**: e1000664. DOI: https://doi.org/10.1371/journal.pcbi.1000664

Goodman DB, Church GM, Kosuri S. 2013. Causes and effects of N-terminal codon bias in bacterial genes. *Science* **342**: 475–479. DOI: https://doi.org/10.1126/science.1241934, PMID: 24072823

Groenenboom MAC, Hogeweg P. 2002. Space and the persistence of male-killing endosymbionts in insect populations. *Proceedings. Biological Sciences* **269**: 2509–2518. DOI: https://doi.org/10.1098/rspb.2002.2197, PMID: 12573064

Groth AC, Fish M, Nusse R, Calos MP. 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* **166**: 1775–1782. DOI: https://doi.org/10.1534/genetics.166.4.1775, PMID: 15126397

Gruber AR, Lorenz R, Bernhart SH, Neuboek R, Hofacker IL. 2008. The Vienna RNA website. *Nucleic Acids Research* **36**: W70–W74. DOI: https://doi.org/10.1093/nar/gkn188

Gu W, Zhou T, Wilke CO, Snel B. 2010. A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLOS Computational Biology* **6**: e1000664. DOI: https://doi.org/10.1371/journal.pcbi.1000664

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**: 696–704. DOI: https://doi.org/10.1080/10635150390235520, PMID: 14530136

Gustafsson C, Govindarajan S, Minshull J. 2004. Codon bias and heterologous protein expression. *Trends in Biotechnology* **22**: 346–353. DOI: https://doi.org/10.1016/j.tibtech.2004.04.006, PMID: 15245907

Hamm CA, Begun DJ, Vo A, Smith CCR, Saelao P, Shaver AO, Jaenike J, Turelli M. 2014. Wolbachia do not live by reproductive manipulation alone: infection polymorphism in *Drosophila suzukii* and *D. subpulchrella*. *Molecular Ecology* **23**: 4871–4885. DOI: https://doi.org/10.1111/mec.12901, PMID: 25156506

Harumoto T, Fukatsu T, Lemaître B. 2018. Common and unique strategies of male killing evolved in two distinct *Drosophila* symbionts. *Proceedings. Biological Sciences* **285**: 20172167. DOI: https://doi.org/10.1098/rspb.2017.2167, PMID: 29563258

Harumoto T, Lemaître B. 2018. Male-killing toxin in a bacterial symbiont of *Drosophila*. *Nature* **557**: 252–255. DOI: https://doi.org/10.1038/s41586-018-0086-2, PMID: 29720654
Hoffmann AA. 1988. Partial cytoplasmic incompatibility between two Australian populations of Drosophila melanogaster. *Entomologia Experimentalis et Applicata* **48**: 61–67. DOI: https://doi.org/10.1111/j.1570-7458.1988.tb02299.x

Hornett E, A, Charlat S, Duploy AMR, Davies N, Roderick GK, Wedell N, Hurst GDD. 2006. Evolution of male-killer suppression in a natural population. *PLOS Biology* **4**: e283. DOI: https://doi.org/10.1371/journal.pbio.0040283, PMID: 16933972

Hornett EA, Duploy AMR, Davies N, Roderick GK, Wedell N, Hurst GDD, Charlat S. 2008. You can’t keep a good parasite down: evolution of a male-killer suppressor uncovers cytoplasmic incompatibility. *Evolution; International Journal of Organic Evolution* **62**: 1238–1263. DOI: https://doi.org/10.1111/j.1558-5646.2008.00353.x, PMID: 18298644

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755. DOI: https://doi.org/10.1093/bioinformatics/17.8.754, PMID: 11524383

Hughes GL, Rasgon JL. 2014. Transinfection: a method to investigate Wolbachia-host interactions and control arthropod-borne disease. *Sect Molecular Biology* **23**: 141–151. DOI: https://doi.org/10.1111/imb.12066, PMID: 24329998

Hunter MS. 1999. The influence of parthenogenesis-inducing Wolbachia on the oviposition behaviour and sex-specific developmental requirements of autoparasitoid wasps. *Journal of Evolutionary Biology* **12**: 735–741. DOI: https://doi.org/10.1046/j.1420-9110.1999.00082.x

Hurst LD. 1997. The incidences and evolution of cytoplasmic male killers. *Proceedings of the Royal Society of London. Series B* **244**: 91–99. DOI: https://doi.org/10.1098/rspb.1991.0056

Hurst GDD, Jiggins FM, Hinrich Graf von der Schulenburg J, Bertrand D, West SA, Goraiacheva I, Zakharov IA, Werren JH, Stouthamer R, Majerus MEN. 1999. Male–killing Wolbachia in two species of insect. *Proceedings of the Royal Society of London. Series B* **266**: 735–740. DOI: https://doi.org/10.1098/rspb.1999.0698

Hurst GD, Jiggins FM. 2000. Male-killing bacteria in insects: mechanisms, incidence, and implications. *Emerging Infectious Diseases* **6**: 329–336. DOI: https://doi.org/10.3201/eid0604.000402, PMID: 1095956

Hurst GD, Johnson AP, Schulenburg JH, Fuyama Y. 2000. Male–killing Wolbachia in *Drosophila*: a temperature-sensitive trait with a threshold bacterial density. *Genetics* **156**: 699–709. DOI: https://doi.org/10.1093/genetics/156.2.699, PMID: 11014817

Hurst GDD, Frost CL. 2015. Reproductive parasitism: maternally inherited symbionts in a biparental world. *Cold Spring Harbor Perspectives in Biology* **7**: a017699. DOI: https://doi.org/10.1101/cshperspect.a017699

Jaenike J, Dyer KA, Reed LK. 2003. Within-population structure of competition and the dynamics of male-killing Wolbachia. *Evolutionary Ecology Research* **5**: 1023–1036.

Jaenike J. 2007. Spontaneous emergence of a new Wolbachia phenotype. *Evolution; International Journal of Organic Evolution* **61**: 2244–2252. DOI: https://doi.org/10.1111/j.1558-5646.2007.00180.x, PMID: 17767593

Jiggins FM, Hurst GD, Majerus ME. 2000. Sex-ratio-distorting Wolbachia causes sex-role reversal in its butterfly host. *Proceedings. Biological Sciences* **267**: 69–73. DOI: https://doi.org/10.1098/rspb.2000.0968, PMID: 10670955

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. *Mammal Prot Metabol* **6**: 538–541. DOI: https://doi.org/10.1038/nature10965, PMID: 22456704

Kudla G, Murray AW, Tollervy D, Plotkin JB. 2009. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**: 255–258. DOI: https://doi.org/10.1126/science.1170160, PMID: 19359587

LePage DP, Metcalf JA, Bordenstein SR, On J, Perlmutter JI, Shropshire JD, Layton EM, Funkhouser-Jones LJ, Freeland SJ, Landweber LF. 2001. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biology* **2**: research0010. DOI: https://doi.org/10.1186/gb-2001-2-4-research0010

Luccchesi JC, Kuroda MI. 2015. Dosage compensation in *Drosophila*. *Cold Spring Harbor Perspectives in Biology* **7**: a019398. DOI: https://doi.org/10.1101/cshperspect.a019398

Majerus TMO, Majerus MEN. 2010. Intergenomic arms races: detection of a nuclear rescue gene of male-killing in a ladybird. *PLOS Pathogens* **6**: e1000987. DOI: https://doi.org/10.1371/journal.ppat.1000987, PMID: 20628578

Marin I, Baker BS. 1998. The evolutionary dynamics of sex determination. *Science* **281**: 1990–1994. DOI: https://doi.org/10.1126/science.281.5385.1990, PMID: 9748152

Masui S, Kuroiwa H, Sasaki T, Inui M, Kuroiwa T, Ishikawa H. 2001. Bacteriophage WO and virus-like particles in Wolbachia, an endosymbiont of arthropods. *Biochemical and Biophysical Research Communications* **283**: 1099–1104. DOI: https://doi.org/10.1006.bbrc.2001.4906, PMID: 11355885

Metcalf JA, Jo M, Bordenstein SR, Jaenike J, Bordenstein SR. 2014. Recent genome reduction of Wolbachia in *Drosophila* recens targets phage WO and narrows candidates for reproductive parasitism. *PeerJ* **2**: e529. DOI: https://doi.org/10.7717/peerj.529
Mitsuhashi W, Ikeda H, Muraji M. 2011. Fifty-year trend towards suppression of Wolbachia-induced male-killing by its butterfly host, Hypolimnas bolina. Journal of Insect Science 11: 92. DOI: https://doi.org/10.1673/031.011.9201, PMID: 21870980

Perlmutter JI, Bordenstein SR, Unckless RL, LePage DP, Metcalf JA, Hill T, Martinez J, Jiggins FM, Bordenstein SR. 2019. The phage gene wmk is a candidate for male killing by a bacterial endosymbiont. PLOS Pathogens 15: e1007936. DOI: https://doi.org/10.1371/journal.ppat.1007936, PMID: 31504075

Perlmutter JI, Meyers JE, Bordenstein SR. 2020. Transgenic testing does not support a role for additional candidate genes in Wolbachia male killing or cytoplasmic incompatibility. MSystems 5: e00658. DOI: https://doi.org/10.1128/mSystems.00658-19, PMID: 31937677

Plotkin JB, Kudla G. 2011. Nonymous but not the same: the causes and consequences of codon bias. Nature Reviews. Genetics 12: 32–42. DOI: https://doi.org/10.1038/nrg2899, PMID: 21102527

Qin H, Wu WB, Comeron JM, Kreitman M, Li WH. 2004. Intragenic spatial patterns of codon usage bias in prokaroytic and eukaryotic genomes. Genetics 168: 2245–2260. DOI: https://doi.org/10.1534/genetics.104.030866, PMID: 15611189

Riparbelli MG, Giordano R, Uyema M, Callaini G. 2012. Wolbachia-mediated male killing is associated with defective chromatin remodeling. PLOS ONE 7: e30045. DOI: https://doi.org/10.1371/journal.pone.0030045, PMID: 22291901

Rodriguez MA, Vermaak D, Bayes JJ, Malik HS. 2007. Species-specific positive selection of the male-specific lethal complex that participates in dosage compensation in Drosophila. PNAS 104: 15412–15417. DOI: https://doi.org/10.1073/pnas.070445104, PMID: 17878295

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574. DOI: https://doi.org/10.1093/bioinformatics/btg180, PMID: 12912839

Saltz R, Kimchi-Sarfaty C, Gottesman MM, Przytycka TM. 2013. Sensitive measurement of single-nucleotide polymorphism-induced changes of RNA conformation: application to disease studies. Nucleic Acids Research 41: 44–53. DOI: https://doi.org/10.1093/nar/gks1009, PMID: 23125360

Sasaki T, Kubo T, Ishikawa H. 2002. Interspecific transfer of Wolbachia between two lepidopteran insects expressing cytoplasmic incompatibility: a Wolbachia variant naturally infecting Cadra cautella causes male killing in Ephestia kuehniella. Genetics 162: 1313–1319. DOI: https://doi.org/10.1093/genetics/162.3.1313, PMID: 12454075

Sauna ZE, Kimchi-Sarfaty C. 2011. Understanding the contribution of synonymous mutations to human disease. Nature Reviews Genetics 12: 683–691. DOI: https://doi.org/10.1038/nr3051

Sharp PM, Averof M, Lloyd AT, Matassi G, Peden JF. 1997. DNA sequence evolution: the sounds of silence. Nature 389: 678–689. DOI: https://doi.org/10.1038/389678a0, PMID: 9258201

Simpson MA, Bordenstein SR, Slatko B. 2018. Microbe Profile: Wolbachia: a sex selector, a viral protector and a therapeutic target to treat filarial nematodes. Microbiology 164: 1345–1347. DOI: https://doi.org/10.1099/mic.0.007024, PMID: 30311871

Southall TD, Elliott DA, Brand AH. 2008. The GAL4 system: A versatile toolkit for gene expression in Drosophila Cold Spring Harbor Protocols 2008: top49. DOI: https://doi.org/10.1101/pdb.top49

Taylor MJ, Bordenstein SR, Slatko B. 2018. Microbe Profile: Wolbachia: a sex selector, a viral protector and a target to treat filarial nematodes. Microbiology 164: 1345–1347. DOI: https://doi.org/10.1099/mic.0.007024, PMID: 30311871

Tuller T, Carmi A, Vestigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y. 2010. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell 141: 344–354. DOI: https://doi.org/10.1016/j.cell.2010.03.031, PMID: 20403328

Unckless RL, Jaenike J. 2012. Maintenance of a male-killing Wolbachia in Drosophila innubila by male-killing dependent and male-killing independent mechanisms. Evolution; International Journal of Organic Evolution 66: 678–689. DOI: https://doi.org/10.1111/j.1558-5646.2011.01485.x, PMID: 22380432

Van Borm S, Wenseleers T, Billen J, Boomsma JJ. 2008. Wolbachia in leafcutter ants: a widespread symbiont that may induce male killing or incompatibility matings. Journal of Evolutionary Biology 14: 805–814. DOI: https://doi.org/10.1111/j.1420-9101.2001.00321.x

Veneti Z, Zabalou S, Papafotiou G, Paraskevopoulou C, Pattas S, Livadaras I, Markakis G, Herren JK, Jaenike J, Bourtis K. 2012. Loss of reproductive parasitism following transfer of male-killing Wolbachia to Drosophila melanogaster and Drosophila simulans. Heredity 109: 306–312. DOI: https://doi.org/10.1038/hdy.2012.43, PMID: 22892635

Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, Brownlie JC, McGraw EA, Martin W, Esser C, Ahmadinejad N, Wiegang C, Madupu R, Beanan MJ, Brinkac LM, Daugherty SC, Durkin AS, Kolonay JF, Nelson WC, Mohamoud Y, Lee P, et al. 2004. Phylgenomics of the reproductive parasite Wolbachia pipientis wMel: a streamlined genome overrun by mobile genetic elements. PLOS Biology 2: E69. DOI: https://doi.org/10.1371/journal.pbio.0020069, PMID: 15024419

Yen JH, Barr AR. 1971. New hypothesis of the cause of cytoplasmic incompatibility in Culex pipiens L. Nature 232: 657–658. DOI: https://doi.org/10.1038/232657a0, PMID: 4937405

Zeh DW, Zeh JA, Bonilla MM. 2005. Wolbachia, sex ratio bias and apparent male killing in the harlequin beetle riding pseudoscorpion. Heredity 95: 41–49. DOI: https://doi.org/10.1038/sj.hdy.6800666, PMID: 15931253