CrossMark
click for updates

Birth of a W sex chromosome by horizontal transfer of *Wolbachia* bacterial symbiont genome

Sébastien Leclercq[a,1], Julien Thézé[a,1], Mohamed Amine Chebbi[a], Isabelle Giraud[a], Bouziane Moumen[a], Lise Ernenwein[a], Pierre Grève[a], Clément Gilbert[a], and Richard Cordaux[a,2]

[a]Laboratoire Ecologie et Biologie des Interactions, Equipe Ecologie Evolution Symbiose, Université de Poitiers, UMR CNRS 7267, 86073 Poitiers Cedex 9, France

Edited by Thomas W. Cline, University of California, Berkeley, CA, and approved October 21, 2016 (received for review June 3, 2016)

Sex determination is a fundamental developmental pathway governing male and female differentiation, with profound implications for morphology, reproductive strategies, and behavior. In animals, sex differences between males and females are generally determined by genetic factors carried by sex chromosomes. Sex chromosomes are remarkably variable in origin and can differ even between closely related species, indicating that transitions occur frequently and independently in different groups of organisms. The evolutionary causes underlying sex chromosome turnover are poorly understood, however. Here we provide evidence indicating that *Wolbachia* bacterial endosymbionts triggered the evolution of new sex chromosomes in the common pillbug *Armadillidium vulgare*. We identified a 3-Mb insert of a feminizing *Wolbachia* genome that was recently transferred into the pillbug nuclear genome. The *Wolbachia* insert shows perfect linkage to the female sex, occurs in a male genetic background (i.e., lacking the ancestral W female sex chromosome), and is hemizygous. Our results support the conclusion that the *Wolbachia* insert is now acting as a female sex-determining region in pillbugs, and that the chromosome carrying the insert is a new W sex chromosome. Thus, bacteria-to-animal horizontal genome transfer represents a remarkable mechanism underpinning the birth of sex chromosomes. We conclude that sex ratio distorters, such as *Wolbachia* endosymbionts, can be powerful agents of evolutionary transitions in sex determination systems in animals.

sex chromosome | horizontal transfer | *Wolbachia* | endosymbiont | isopod crustacean

In many animals, sex determination is controlled by sex chromosomes. Sex chromosomes are heterozygous in males in XY/XX systems (male heterogamety) and in females in ZW/ZZ systems (female heterogamety) (1–4). Despite their involvement in a fundamental developmental process that might be expected to be highly conserved, sex determination systems are highly variable and generally show a poor fit to phylogeny (5–8). The remarkable lability of sex determination systems implies a complex evolutionary history of transitions between systems and the recurrent evolution of new sex chromosomes.

New sex chromosomes arise from ordinary autosomes that acquire a sex-determining region, which is characterized by sex-specific inheritance and is often hemizygous (i.e., a haploid region in an otherwise diploid individual) in the heterogametic sex (9). The evolutionary forces driving the evolution of new sex chromosomes and the mechanisms underlying transitions between sex determination systems are poorly understood, however (2, 3).

It has been proposed that sex ratio selection may be the major evolutionary force driving turnovers of sex determination systems (3, 10, 11). Sex ratio selection occurs to restore a balanced sex ratio (the usual evolutionarily stable strategy) in a population in which the sex ratio is biased toward one sex. Sex ratio biases are often caused by sex ratio distorters, selfish genetic elements with sex-biased inheritance that enhance their own transmission by biasing sex ratios to their advantage. Examples of these sex ratio distorters include meiotically driven sex chromosomes and cytoplasmic elements, such as organelles and endosymbiotic microorganisms (12–15). Their presence generates genetic conflicts with most other genes in a genome, which are biparentally inherited. Ensuing genetic conflicts may lead to the evolution of nuclear suppressors restoring balanced sex ratios and turnovers in sex determination systems (2, 3). Here we extend the concept that sex ratio distorters may drive the evolution of sex-determination systems through direct incorporation in the nuclear genome as sex-determining factors.

The common pillbug *Armadillidium vulgare* (Crustacea, Isopoda) has a female heterogametic system of sex determination with ZW females and ZZ males (15–17); however, many *A. vulgare* females produce female-biased progenies owing to intracellular, maternally inherited *Wolbachia* bacterial endosymbionts (14–17). These cytoplasmic sex ratio distorters induce the feminization of genetic males into phenotypic females (18, 19). ZZ genetic male embryos carrying *Wolbachia* develop as viable and fertile phenotypic females, which in turn produce *Wolbachia*-infected, female-biased progenies. It is predicted by models, and has been confirmed empirically, that ZW genetic females will be eliminated from pillbug lines infected by *Wolbachia*, causing the loss of the W sex chromosome (20). This is a perfect example of transition from chromosomal to cytoplasmic sex determination (15–17). All individuals are then ZZ genetic males; those inheriting *Wolbachia* develop as females, whereas uninfected embryos develop as males.

Interestingly, however, another sex determination factor exists in this species, with many female-biased progenies also produced in some lines lacking *Wolbachia* (21). Females from these lines are thought to be ZZ genetic males converted into females by an unknown feminizing agent termed the "*f* element." Based on an

Significance

Sex determination is an evolutionarily ancient, key developmental pathway governing sexual differentiation in animals. Sex determination systems are remarkably variable between species or groups of species, however, and the evolutionary forces underlying transitions between these systems are poorly understood. Here we provide evidence indicating that the genome of *Wolbachia* bacterial endosymbionts was horizontally transferred into a chromosome of the common pillbug *Armadillidium vulgare*, which resulted in this chromosome evolving as a new female (W) sex chromosome. This represents a remarkable mechanism underpinning the birth of sex chromosomes. We conclude that bacteria can be powerful agents of evolutionary transitions in sex determination systems in animals.

Author contributions: R.C. designed research; S.L., J.T., M.A.C., I.G., B.M., and L.E. performed research; P.G. contributed new reagents/analytic tools; S.L., J.T., M.A.C., I.G., B.M., P.G., C.G., and R.C. analyzed data; and R.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequencing data generated in the study have been deposited at DNA Data Bank of Japan/European Nucleotide Archive/GenBank (accession no. LYUU00000000). The version described in the paper is version LYUU01000000.

[1]S.L. and J.T. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: richard.cordaux@univ-poitiers.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1608979113/-/DCSupplemental.

indirect physiological test, it has been suggested that this *f* element is widely distributed in field populations of *A. vulgare*, and its frequency is consistent with a recent origin (22). In 1984, Legrand and Juchault (21) hypothesized that the *f* element may be a fragment of a *Wolbachia* genome integrated in the pillbug nuclear genome and carrying feminization information. In the present study, we tested this more than 30-y-old hypothesis.

Results and Discussion

To identify the *f* element, we sequenced the female genome from an *f*-carrying pillbug line derived from wild animals sampled in Denmark. Mapping of 622 million sequencing reads onto 16 *Wolbachia* genome sequences revealed that 0.11% of the reads showed high similarity to *Wolbachia*. These reads were assembled into nine scaffolds spanning 3.13 Mb, including one scaffold of 2.8 Mb and eight scaffolds of 10–106 kb (Table 1). Phylogenetic comparison with 19 representative *Wolbachia* strains at six genetic loci (3,038 nucleotides) revealed that the *Wolbachia* scaffolds are most closely related to the feminizing *Wolbachia* strains of *A. vulgare* (*w*VulC) and *A. nasatum* (*w*Nas), within the Oni clade of the B supergroup of *Wolbachia* diversity (23) (Fig. 1). A comparison of 839 orthologous genes (841,039 nucleotides) confirmed that the *Wolbachia* scaffolds are nearly identical to *Wolbachia* *w*VulC at the genome-wide level (99.7% nucleotide identity). This extremely low level of sequence divergence (only 3.5 substitutions per gene on average) indicates that the *Wolbachia* scaffolds and *w*VulC diverged from one another only very recently.

We assessed whether the *Wolbachia* scaffolds belong to a cytoplasmic *Wolbachia* genome or to nuclear insert(s) using transmission electron microscopy, which indicated the absence of *Wolbachia* and any other endosymbionts in the cytoplasm of oocytes of *A. vulgare* females carrying the *f* element (Fig. 2*A*), in agreement with previous observations (21). Furthermore, our sequencing reads covered only 82.6% of the 1.66-Mb *Wolbachia* *w*VulC genome (Fig. 3). Inspection of the nonmapped regions indicated that 331 out of 1,888 *w*VulC genes are absent from ($n = 243$) or severely truncated in ($n = 88$) the *Wolbachia*-derived scaffolds from the *f*-carrying female (Table S1). Importantly, 85 of these 331 missing or partial genes (26%) belong to the *Wolbachia* core genome, i.e., they are crucial to *Wolbachia* cell viability (Table S1). Because the *Wolbachia* scaffolds lack these genes, we conclude that they do not belong to a cytoplasmic *Wolbachia* genome and instead have a nuclear origin. Nuclear *Wolbachia* inserts have been previously reported in the genomes of other arthropods and nematodes (24–26).

To test whether the reconstructed scaffolds belong to one or more *Wolbachia* inserts integrated in the pillbug genome, and whether one of these is the *f* element, we performed PCR analyses specifically targeting each of the nine *Wolbachia* scaffolds (between two and eight markers per scaffold). Analysis of a panel of males and females from an *f*-carrying pillbug line showed PCR amplification in

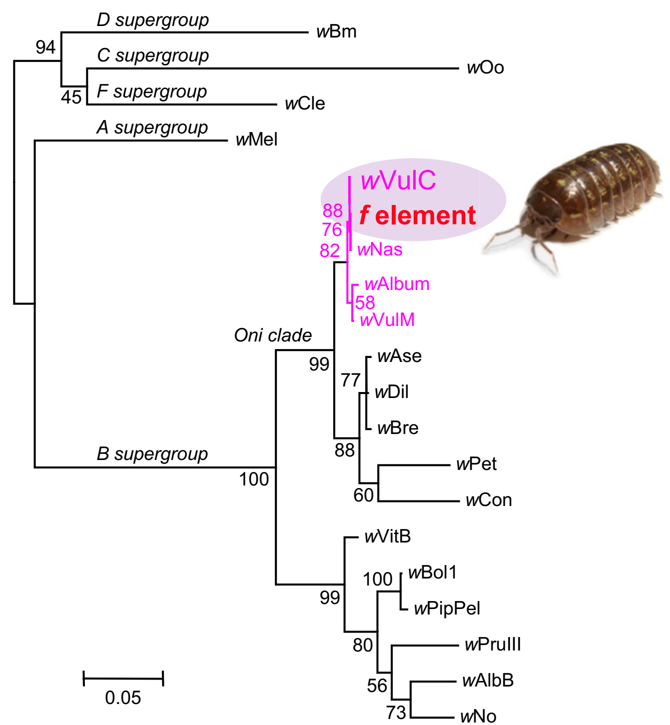

Fig. 1. Close relationship between the *f* element and feminizing *Wolbachia* *w*VulC from the common pillbug *A. vulgare*. The maximum likelihood phylogenetic tree of the *f* element and representative *Wolbachia* strains is shown. Bootstrap scores are shown as percentages. The scale bar shows estimated nucleotide divergence in nucleotide substitutions per site. The *f* element from *A. vulgare* (bold red) is nested within a clade of feminizing *Wolbachia* strains (pink) and is closely related to strain *w*VulC from the common pillbug *A. vulgare* (purple oval and photograph).

all females and in none of the males for all scaffolds except scaffold 9 (Table 1). Cosegregation of scaffolds 1–8 with the female sex indicated that they are genetically linked and most likely belong to a single *Wolbachia* insert. Scaffold 9 is present in both males and females and thus represents an independent event of horizontal transfer from *Wolbachia* to the pillbug genome that cannot be the *f* element. In contrast, the *Wolbachia* insert corresponding to scaffolds 1–8 is linked to the female sex, as expected for the *f* element. Female sex linkage was independently confirmed by a pedigree analysis spanning six generations (118 descendants screened), which showed perfect maternal inheritance of the *Wolbachia*

Table 1. Comparison of nine *Wolbachia* scaffolds integrated in the female genome of *A. vulgare* carrying the *f* element

Sequence name	Size, bp	PCR amplification	Mean sequencing depth	*f* element
Scaffold 1*	2,798,100	All ♀, no ♂	21.7×	Yes
Scaffold 2	106,076	All ♀, no ♂	20.1×	Yes
Scaffold 3	44,965	All ♀, no ♂	23.6×	Yes
Scaffold 4	64,497	All ♀, no ♂	22.6×	Yes
Scaffold 5	67,565	All ♀, no ♂	20.4×	Yes
Scaffold 6	13,410	All ♀, no ♂	20.2×	Yes
Scaffold 7	13,655	All ♀, no ♂	21.6×	Yes
Scaffold 8	10,104	All ♀, no ♂	21.2×	Yes
Scaffold 9	10,554	Some ♀, some ♂	40.3×	No
2,087 *A. vulgare* scaffolds ≥5 kb	13,401,219	Not tested	41.5×	No

*The *f* element starts at nucleotide position 95,380. Nucleotide positions 1–95,379 belong to the *A. vulgare* flanking sequence.

EVOLUTION

PNAS PNAS

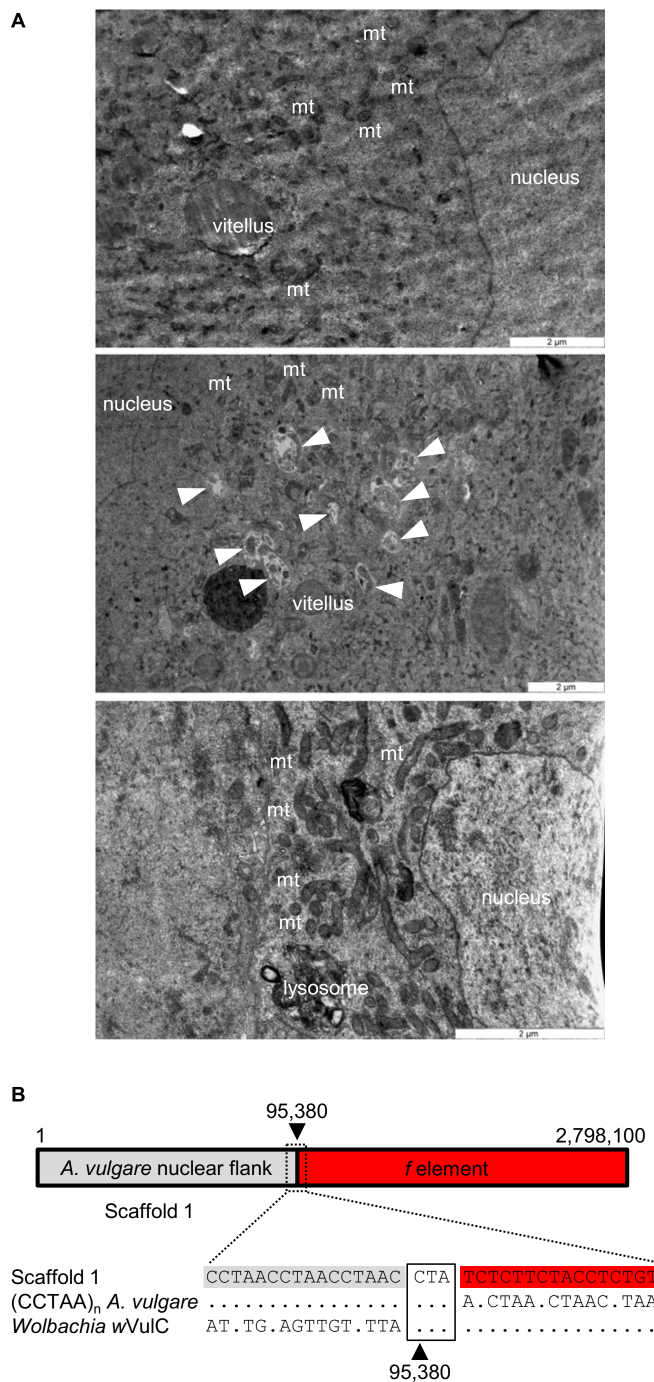

Fig. 2. Nuclear localization of the *f* element. (*A*) Transmission electron microscopy of oocytes of *A. vulgare* females carrying the *f* element (*Top*), infected by *Wolbachia w*VulC (*Middle*), or lacking both the *f* element and *Wolbachia* (*Bottom*). *Wolbachia* bacterial cells in the cytoplasm of oocytes are indicated by white arrows. Oocytes from *f* element-carrying female lack cytoplasmic *Wolbachia* cells. mt, mitochondria. (Scale bars: 2 μm.) (*B*) Schematic representation (not to scale) of scaffold 1 from the *f* element (*Top*) and zoom-in region encompassing junction between the *f* element (red) and *A. vulgare* nuclear sequence (gray) at nucleotide resolution (*Bottom*). Nucleotide coordinates are shown. The 3-bp microhomology present on both the flanking *A. vulgare* $(CCTAA)_n$ microsatellite and *Wolbachia w*VulC sequence is boxed. Dots indicate nucleotides identical to the scaffold 1 sequence.

insert (Fig. 4). Furthermore, scaffolds 1–8 showed a mean sequencing depth of ∼21×, one-half of the sequencing depth of ∼41× observed for the diploid *A. vulgare* genome and autosomal scaffold 9 (Table 1). This indicates that the female-linked *Wolbachia* insert is hemizygous, as expected for the sex-determining, W chromosome-specific region of the genome.

These findings raise the question of whether the *Wolbachia* insert was integrated in the female sex-determining region of the native W sex chromosome of the pillbug genome, or was integrated in another genomic locus that has since become a new sex-determining region. To distinguish between these two hypotheses, we performed genetic crosses spanning three generations (*Materials and Methods*). We predicted that F2 progenies should exclusively consist of females if the *Wolbachia* insert were linked to the native female sex-determining region of the pillbug genome, or up to 50% males if the *Wolbachia* insert occurred in a ZZ genetic male background (i.e., lacking the native female sex-determining region). We found that all 25 F2 progenies (939 individuals) were composed of >21% males, thus verifying the second prediction (Table S2). These results provide direct evidence that the *Wolbachia* insert was integrated in a genetic background lacking the female sex-determining region of the native W sex chromosome. Therefore, the *Wolbachia* insert does not result from integration in the native W sex chromosome. Instead, female linkage of the *Wolbachia* insert, together with the known feminizing effect of the cytoplasmic *Wolbachia* ancestor, strongly support the view that the *Wolbachia* insert acts as a new female sex-determining region of the genome. We acknowledge that a definitive demonstration of the feminizing nature of the *Wolbachia* insert will require functional validation; however, owing to the large number of potentially functional genes that it contains, at present we are unable to perform such experiments.

To investigate the genomic context of the *Wolbachia* insert, we characterized a 95-kb-long region at the 5′ end of scaffold 1 that showed no similarity to *Wolbachia*. Five loci indicate that this flanking sequence is derived from the pillbug genome; three of these loci show the greatest similarity to homologs found in animal genomes, and the other two loci are fragments of viral genomes of the crustacean-borne Nimaviridae family endogenized into the host genome (27) (Table S3). This flanking genomic region of pillbug origin provides direct evidence for the nuclear localization of the *Wolbachia* insert. Furthermore, inspection of the sequence at the junction between the pillbug genome and the *Wolbachia* insert showed that integration occurred in a $(CCTAA)_n$ microsatellite. The integration site shows a 3-bp microhomology (Fig. 2*B*), suggesting that the *Wolbachia* sequence was likely integrated in the pillbug genome as part of the repair of a DNA double-strand break (28, 29).

Nearly 83% of the *Wolbachia* genome is currently found in the *Wolbachia* insert, but up to the entire *Wolbachia* genome may have been transferred, with subsequent deletions of small genomic regions. Consistent with the latter hypothesis, the 331 *w*VulC genes missing or severely truncated in the *Wolbachia* insert are clustered in 55 regions of 1–40 contiguous genes in the *w*VulC genome, with six contiguous genes per cluster on average. Nevertheless, we cannot formally exclude the possibility that a chunk of the *Wolbachia* genome was lost on integration. The original transfer of 83% (conservatively assuming all 331 genes lost on integration) to 100% (assuming no genes lost on integration) of the *Wolbachia* genome means that the original size of the *Wolbachia* insert on integration in the pillbug genome was 1.37–1.66 Mb (using *w*VulC as a reference). Thus, the *Wolbachia* insert roughly doubled in size (3.02 Mb) since its integration owing to numerous duplications of varying sizes (Fig. 3). We identified 88 duplications (excluding mobile elements) ranging in size from 1.5 to 170 kb, which collectively contributed 1.46 Mb of additional sequence to the *Wolbachia* insert (Fig. 3), although *w*VulC and the *Wolbachia* insert are nearly identical at the nucleotide level (99.7% nucleotide identity). Thus, genomic rearrangements mediated by intrachromosomal recombination appear to be a major signature of the early molecular evolution of the *Wolbachia* insert.

Conclusion

Taken together, our results validate the hypothesis proposed by Legrand and Juchault in 1984 (21), demonstrating that the *f* element is a large piece of a feminizing *Wolbachia* genome that has

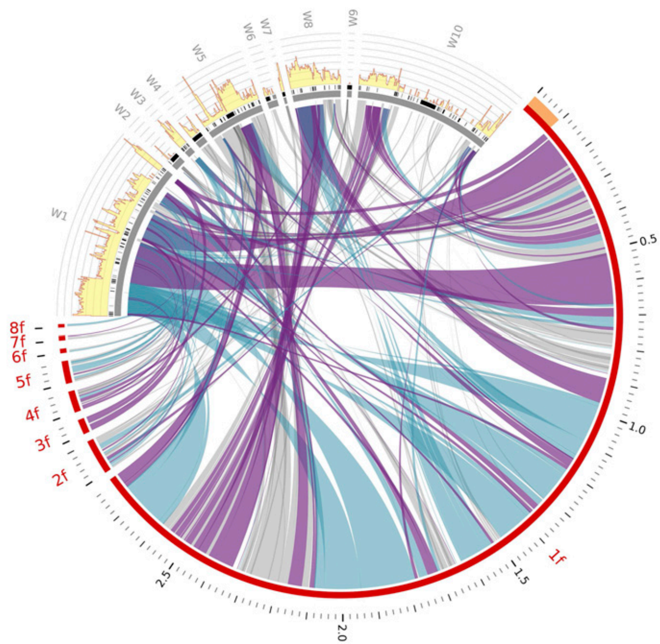

Fig. 3. Genomic comparison of the *f* element and *Wolbachia* *w*VulC. Scaffolds of the *f* element (1*f*–8*f*) and *Wolbachia* *w*VulC (W1–W10) are shown in red and gray, respectively (scale in Mb). Purple and blue links indicate ≥1.5-kb regions duplicated in the *f* element compared with *w*VulC, in forward and reverse orientation, respectively. Gray links indicate unique regions and mobile elements. Black bars indicate annotated mobile elements on *w*VulC scaffolds. The yellow histogram shows mapping coverage of sequencing reads onto *w*VulC scaffolds. Gray arcs represent 20× coverage scale. The orange region is the pillbug nuclear region.

recently been transferred horizontally to the pillbug nuclear genome. The *f* element was not integrated in the native W sex chromosome. Perfect female sex linkage and hemizygosity, together with the known feminizing effect of the cytoplasmic *Wolbachia* ancestor, strongly support the view that the *f* element is acting as the female sex-determining region in *f*-carrying pillbug lines. In any event, we show that the pillbug chromosome in which the *f* element is inserted has effectively become a new W sex chromosome. Our results thus support an evolutionary scenario in which *Wolbachia* endosymbionts triggered a turnover of sex chromosomes in *A. vulgare* (15–17, 20), first by inducing the loss of the native W sex chromosome under cytoplasmic sex determination and then by introducing a new sex-determining region elsewhere in the nuclear genome (Fig. 5). This scenario implies that the *f* element is a sex-determining region of bacterial origin in the genome of an animal species, and that bacteria-to-animal horizontal genome transfer can be viewed as a novel mechanism for the birth of sex chromosomes. The *f* element is remarkable in that it represents an extremely recent case of horizontal transfer with major biological consequences (individual development as female, pending functional validation), which allowed us to trace back the entire evolutionary history of the transfer, including donor (feminizing *Wolbachia* bacteria closely related to *w*VulC), mechanism (integration of a nearly complete genome by microhomology-mediated recombination), molecular evolution (genomic expansion by intrachromosomal recombination), and likely evolutionary significance (recruitment as a sex-determining region). We conclude that sex ratio distorters such as *Wolbachia* bacterial endosymbionts can be powerful agents of evolutionary transitions in sex determination systems in animals.

Materials and Methods

Genome Sequencing and Assembly. We extracted total genomic DNA from four *f*-carrying *A. vulgare* sisters using the Qiagen DNeasy Blood and Tissue Kit, according to the protocol for animal tissues (3 h of incubation in proteinase K at 56 °C and 30 min of RNase treatment at 37 °C). These females were from our laboratory line WXf-1543 derived from wild animals sampled in Helsingor, Denmark in 1989. One paired-end (~250-bp inserts) and three mate-pair (~4.2-kb inserts) sequencing libraries were prepared (each from a single female) and sequenced as 100-bp reads on the Illumina HiSeq2000 platform. We obtained 783,979,240 paired-end reads (78.4 Gb) and 216,542,354 mate-pair reads (21.7 Gb). Read quality was checked with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/), and low-quality reads were filtered out with Fastx-toolkit (hannonlab.cshl.edu/fastx_toolkit/index.html), leaving 622,379,086 paired-end reads (56.0 Gb) and 77,677,246 mate-pair reads (5.5 Gb). Processed reads were mapped onto 16 *Wolbachia* genomes (Table S4) using Bowtie2 (30). Mapped reads were assembled using Velvet version 1.2.10 (31) (*k*-mer size of 55), SOAPdenovo version 1.05 (32) (*k*-mer size of 61), and Allpaths-LG version R43987 (33). Each assembly was optimized by scaffolding preassembled contigs with SSpace version 2.0 (34), and GapFiller version 1.10 (35) was then used to minimize undetermined nucleotides. The three assemblies were then merged with Minimus2 (36), resulting in 46 *Wolbachia*-like scaffolds totaling 1,740,353 bp.

Sequencing depth was heterogeneous within and between scaffolds, with blocks of homogeneous depth (roughly multiples of ~20× depth) separated by sudden shifts coinciding with misassemblies, suggesting the occurrence of multiple duplicated regions in the assembly. Therefore, scaffolds were broken into blocks of homogeneous sequencing depth, and misassemblies were resolved by iterative mapping using Geneious version 7.1 (37) and manual curation. The final assembly was composed of nine *Wolbachia*-like scaffolds totaling 3,128,926 bp (Table 1). Automated annotation was performed using Prokka version 1.9 (38).

In parallel, sequencing reads were used to assemble the *A. vulgare* genome, using SOAPdenovo version 1.05. The genome (*k*-mer size of 49) was composed of 3,549,649 scaffolds and contigs totaling 1.456 Gb. To remove contaminant sequences, the 3,768 scaffolds ≥5 kb in length were used for similarity searches by blastx version 2.2.30+ (39) against the nonredundant protein sequences of the National Center for Biotechnology Information (NCBI) database (as of September 2015). The 2,087 scaffolds assigned to eukaryotes were considered to belong to the

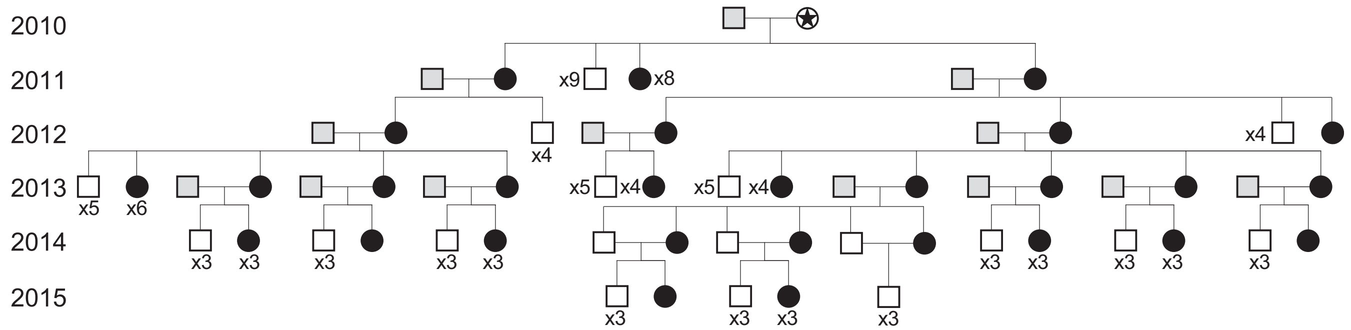

Fig. 4. Perfect maternal inheritance of the *f* element. A pedigree of an *f* element-carrying *A. vulgare* family spanning six generations (one per year; years on the left) is shown. Males are shown as squares; females, as circles. Individuals carrying or lacking the *f* element are shown in black and white, respectively. The founding female carrying the *f* element is indicated by a star, and males from other lines (lacking both the *f* element and *Wolbachia*) used for crossings are shown in gray. The presence/absence of the *f* element and *Wolbachia* was confirmed by PCR in all individuals except the two founders.

PNAS PNAS PNAS

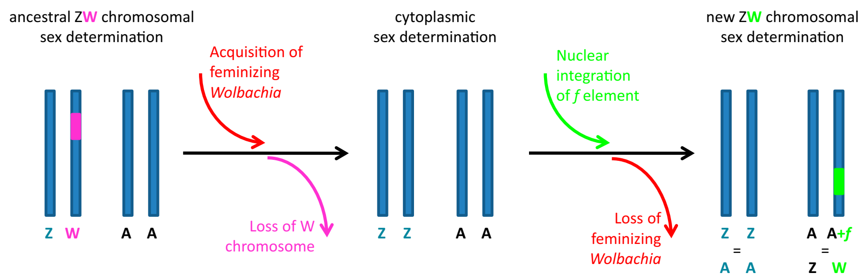

Fig. 5. Evolutionary scenario for *Wolbachia*-mediated turnover of sex chromosomes in the common pillbug *A. vulgare*. In the background of an ancestral ZW heterogametic system (*Left*), acquisition of feminizing *Wolbachia* endosymbionts resulted in loss of the ancestral W sex chromosome and evolution of cytoplasmic sex determination (*Middle*). Subsequent integration of a *Wolbachia* genome piece carrying feminization information (*f* element) in an autosome of the pillbug genome (A) and secondary loss of *Wolbachia* endosymbionts resulted in the evolution of a new ZW heterogametic system (*Right*).

A. vulgare genome per se. (The other scaffolds were assigned to prokaryotes or showed no significant hit.) Illumina paired-end processed reads were mapped onto the nine *Wolbachia*-like scaffolds and the *A. vulgare* genome using Bowtie2. The mean sequencing depth per nucleotide of the *Wolbachia*-like scaffolds and the 2,087 *A vulgare* scaffolds ≥5 kb in length was determined using fitGCP (40).

Molecular Tests. We analyzed 26 molecular markers targeting the nine *Wolbachia* scaffolds (two to eight markers per scaffold) by PCR (Table S5). PCR reactions were carried out in 25 μL with 5 μL of buffer 5×, 0.5 μL of dNTPs (2.15 mM), 0.7 μL of each primer (10 μM), 0.25 μL of Taq polymerase 5 u/μL, and 1 μL of DNA. PCRs were conducted using the following temperature cycling: initial denaturation at 94 °C for 3 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 47–60 °C (depending on primer set) for 30 s, and elongation at 72 °C for 1 min, ending with a 10-min elongation step at 72 °C. Amplification patterns in a panel of six males and six females of the WXf-1543 line indicated that scaffold 9 does not belong to the *f* element; thus, scaffold 9 was removed from all subsequent analyses, and scaffolds 1–8 are hereinafter referred to as the *f* element.

To distinguish between the presence of the *f* element and *Wolbachia* infection, we used a combination of four molecular markers: *wsp*, *recR*, *dnaA*, and *ftsZ* (Table S5). Whereas all four of these markers are present in the *Wolbachia* genome, only two are present in the *f* element (*dnaA* and *ftsZ*). Thus, different amplification patterns are obtained for individuals with the *f* element (*wsp*−, *recR*−, *dnaA*+, *ftsZ*+), *Wolbachia* (*wsp*+, *recR*+, *dnaA*+, *ftsZ*+), or none (*wsp*−, *recR*−, *dnaA*−, *ftsZ*−). In addition, we designed a marker (Jtel) spanning the junction between the *f* element and the pillbug genome (Table S5) as an alternative assay for *f* element detection and to distinguish between individuals infected only by *Wolbachia* and those with both *Wolbachia* and the *f* element.

Phylogenetic Analyses. Nucleotide sequences of six genes (*cox*A, *fbp*A, *ftsZ*, *gat*B, *gro*E, and *hcp*A) widely used for *Wolbachia* phylogenetic studies (41, 42) and previously sequenced in 19 representative *Wolbachia* strains (Table S6) were used for similarity searches by blastn against the *f* element. Some of these genes were duplicated in the *f* element (Table S6), but all paralogous copies had identical sequences. For each taxon, the six loci were aligned and concatenated (total alignment length, 3,038 nucleotides). MEGA6 (43) was used to identify the best substitution model (TN93+G+I), and a maximum likelihood phylogeny was reconstructed using PhyML 3.0 (44), with 1,000 bootstrap replicates.

Comparisons with *Wolbachia* *w*VulC. Amino acid sequences of the 3,173 genes from the *f* element were used for similarity searches by blastp version 2.2.30+ against the 1,888 genes of the *Wolbachia* *w*VulC genome (GenBank accession no. ALWU00000000). The 839 best hits matching >100-aa-long *w*VulC proteins and aligning to >90% of their length with proteins of the *f* element were retained. Nucleotide sequences of the 839 genes from *w*VulC and the *f* element were aligned (total alignment length, 841,039 nucleotides), and mismatches were counted using Emboss Stretcher (45). We found that 633 of the 839 sequence pairs were identical, and that 796 of the 839 sequence pairs exhibited >99% nucleotide identity.

Illumina paired-end processed reads (see above) were mapped onto the *Wolbachia* *w*VulC genome using Bowtie2. For each of the 1,888 wVulC genes, we calculated the fraction of nucleotides mapped by *f* element reads. We found that 1,492 *w*VulC genes were fully mapped (i.e., present in the *f* element), 243 genes were not mapped at all (i.e., absent in the *f* element), and 153 genes were partially mapped. Of the latter, the 88 genes with <80% mapped nucleotides were considered severely truncated and likely nonfunctional in the *f* element. *Wolbachia* core genes absent or nonfunctional in the *f* element were identified using similarity searches by blastp against the 654 genes of the *Wolbachia* core genome (46). Of the 331 genes lacking or severely truncated in the *f* element, 85 belong to the *Wolbachia* core genome (based on >75% amino acid similarity and >80% alignment length to a *Wolbachia* core gene, using *w*PipPel as a reference).

Genomic duplications in the *f* element were identified by similarity searches against *w*VulC scaffolds using blastn. Hits covered on >80% of their length by another hit on the *f* element were filtered out using an in-house perl script, as were hits <1,500 bp in length or showing <70% nucleotide similarity with the *w*VulC sequence. Duplications matching mobile elements (regions annotated as IS transposable elements, group II introns, and prophages in *w*VulC) on >90% of their length were also filtered out to exclude transposition events from the count of genomic duplications. Results were visualized using CIRCOS 0.69 (47).

Transmission Electron Microscopy. Observations were performed on ovaries of two adult females from each of three *A. vulgare* laboratory lines: WXf-1543 (carrying the *f* element), WXw-1651 (infected by *Wolbachia* *w*VulC), and BF-2869 (lacking both the *f* element and *Wolbachia*). For each female, total genomic DNA was extracted from one ovary, and the presence or absence of *Wolbachia* and the *f* element was verified by PCR as described above. The second ovary of each female was fixed in a solution containing equal volumes of glutaraldehyde 3% (wt/vol), sodium cacodylate 0.1 M, and NaCl 1% at pH 7.35 for 2 h at 4 °C. Ovaries were postfixed in a solution containing equal volumes of osmium tetroxide 1.3% (wt/vol), sodium cacodylate 0.1 M, and NaCl 1.8% (wt/vol). After dehydration through an acetone series, samples were embedded in Epon resin. Ultrathin (70 nm) sections were obtained with an Ultracut ultramicrotome (Leica) and a diamond knife (Diatome). Sections were transferred to carbon-coated films on 1-mm copper whole grids and contrasted with 2% uranyl acetate and lead citrate. Observations were performed with a JEOL JEM 1010 transmission electron microscope equipped with an Olympus Quemesa digital camera.

Pedigree Analysis. We generated a pedigree spanning six generations (*A. vulgare* has a generation time of 1 y), starting in 2011 with a cross between two founders, a 1-y-old male and a 1-y-old female carrying the *f* element, the progeny of which constituted the first generation of line WXf-1543. In 2012, 2013, and 2014, between two and seven females were selected from the previous generation and crossed with males from our laboratory lines BF-2794 or BF-2869. In 2015, three females selected from the 2014 generation were crossed with their brothers. At each generation, total genomic DNA was extracted from the progenitors (except the two founders) and from randomly selected males and females of their progenies. Presence or absence of the *f* element and *Wolbachia* was tested by PCR using markers *wsp*, *recR*, *dnaA*, *ftsZ*, and Jtel (as described above) in 130 individuals of the pedigree, including 118 descendants (62 males and 56 females) and 12 males from lines BF-2794 or BF-2869. As expected, none of the 12 males from lines BF-2794 or BF-2869 carried the *f* element or *Wolbachia*.

Genetic Crosses. To test whether the *f* element was integrated in the female sex-determining region of the native W chromosome or elsewhere in the genome and has since become the female sex-determining region, we first generated *A. vulgare* males homozygous for the W sex chromosome (i.e., WW males). We took advantage of the existence of a dominant masculinizing locus (M) in our laboratory line BFog-2812, which is epistatic to the W sex chromosome and the *f* element (48). This makes it possible to generate WW,Mm individuals, which differentiate as males (48). We crossed WW,Mm males with females carrying the *f* element (from our laboratory line WXf-1538) and obtained eight F1 progenies. Total genomic DNA was extracted from one leg of 80 F1 females, and PCR was performed as described above to identify those carrying the *f* element. We then crossed F1 females carrying the *f* element with genetic males (ZZ,mm) from our laboratory line BF-2787 (which lacks both the *f* element and *Wolbachia*). We obtained 25 F2 progenies, whose sex ratios allowed us to discriminate whether the *f* element was integrated in the native W chromosome or elsewhere in the genome, as explained below.

PNAS

If the *f* element is integrated in the native W chromosome (W^f), then the genotype of F0 females carrying the *f* element is ZW^f,mm. Thus, F2 progenies are expected to consist exclusively of females, given the genotypes of the individuals used for the two genetic crosses:

First cross: WW,Mm (♂) × ZW^f,mm (♀) → WZ,Mm (♂) + WW^f,Mm (♂) + WZ,mm (♀) + $\underline{WW^f\text{,mm (♀)}}$

Second cross: ZZ,mm (♂) × $\underline{WW^f\text{,mm (♀)}}$ → ZW,mm (♀) + ZW^f,mm (♀).

If the *f* element is integrated elsewhere in the genome (i.e., in a ZZ genetic background), then the genotype of F0 females carrying the *f* element is ZZ,mm,*f*+*f*− (because the *f* element is hemizygous). A hallmark of this hypothesis is that F2 progenies are expected to contain up to 50% males, given the genotypes of the individuals used for the two genetic crosses:

First cross: WW,Mm,*f*−*f*− (♂) × ZZ,mm,*f*+*f*− (♀) → WZ,Mm,*f*+*f*− (♂) + WZ, Mm,*f*−*f*− (♂) + <u>WZ,mm,*f*+*f*− (♀)</u> + WZ,mm,*f*−*f*− (♀)

Second cross: ZZ,mm,*f*−*f*− (♂) × <u>WZ,mm,*f*+*f*− (♀)</u> → ZW,mm,*f*+*f*− (♀) + ZW,mm,*f*−*f*− (♀) + ZZ,mm,*f*+*f*− (♀) + ZZ,mm,*f*−*f*− (♂).

Characterization of the *f* Element Flanking Region. We identified the integration site of the *f* element in the pillbug genome at coordinate 95,380 in scaffold 1. The 38 ORFs >400 bp long in the genomic region upstream of the integration site (i.e., 5′ end of scaffold 1) were translated and used for similarity searches using blastp against the nonredundant protein sequences of the NCBI database (as of January 2016). The 10 ORFs showing significant hits (e-value $<10^{-5}$) corresponded to six genomic loci. We used the taxonomy of the blastp hits [eukaryote, bilateria (metazoan taxon to which the pillbug belongs), virus, or prokaryote] to infer the origin of the six genomic loci. Whereas one locus yielded ambiguous results, the other five loci supported the conclusion that the first 95,379 bp at the 5′ end of scaffold 1 are of pillbug origin (Table S3).

The immediate pillbug sequence flanking the *f* element is a $(CCTAA)_{92}$ microsatellite (coordinates 94,919–95,378 in scaffold 1), followed by CCTA (coordinates 95,379–95,382). This sequence perfectly matches the first four nucleotides of the microsatellite motif CCTAA and indicates that the *f* element is integrated in a $(CCTAA)_n$ microsatellite of the pillbug genome. Interestingly, a similarity search by blastn using *Wolbachia w*VulC as a query against scaffold 1 unambiguously indicated that the *f* element starts at coordinate 95,380 (Fig. 2*B*). Thus, 3 bp at the *f* element integration site (CTA at coordinates 95,380–95,382) are shared by the pillbug genome and the *f* element, defining a microhomology.

ACKNOWLEDGMENTS. We thank S. Beltran-Bech, J. Peccoud, and T. Rigaud for discussions and comments on the manuscript, and M. Raimond for technical help with transmission electron microscopy. We also thank the editor and two anonymous reviewers for their constructive comments on previous versions of the manuscript. This work was funded by European Research Council Starting Grant 260729 (EndoSexDet) and Agence Nationale de la Recherche Grant ANR-15-CE32-0006-01 (CytoSexDet) (to R.C.), the 2015–2020 State-Region Planning Contract and European Regional Development Fund, and intramural funds from the Centre National de la Recherche Scientifique and the University of Poitiers.

1. Matson CK, Zarkower D (2012) Sex and the singular DM domain: Insights into sexual regulation, evolution and plasticity. *Nat Rev Genet* 13(3):163–174.
2. Bachtrog D, et al.; Tree of Sex Consortium (2014) Sex determination: Why so many ways of doing it? *PLoS Biol* 12(7):e1001899.
3. Beukeboom LW, Perrin N (2014) *The Evolution of Sex Determination* (Oxford Univ Press, Oxford, UK), p 222.
4. Bachtrog D, et al. (2011) Are all sex chromosomes created equal? *Trends Genet* 27(9):350–357.
5. Sarre SD, Ezaz T, Georges A (2011) Transitions between sex-determining systems in reptiles and amphibians. *Annu Rev Genomics Hum Genet* 12:391–406.
6. Kikuchi K, Hamaguchi S (2013) Novel sex-determining genes in fish and sex chromosome evolution. *Dev Dyn* 242(4):339–353.
7. Vicoso B, Bachtrog D (2015) Numerous transitions of sex chromosomes in Diptera. *PLoS Biol* 13(4):e1002078.
8. Holleley CE, et al. (2015) Sex reversal triggers the rapid transition from genetic to temperature-dependent sex. *Nature* 523(7558):79–82.
9. Bergero R, Charlesworth D (2009) The evolution of restricted recombination in sex chromosomes. *Trends Ecol Evol* 24(2):94–102.
10. Werren JH, Beukeboom LW (1998) Sex determination, sex ratios, and genetic conflict. *Annu Rev Ecol Syst* 29:233–261.
11. Mank JE, Hosken DJ, Wedell N (2014) Conflict on the sex chromosomes: Cause, effect, and complexity. *Cold Spring Harb Perspect Biol* 6(12):a017715.
12. Jaenike J (2001) Sex chromosome meiotic drive. *Annu Rev Ecol Syst* 32:25–49.
13. Chase CD (2007) Cytoplasmic male sterility: A window to the world of plant mitochondrial-nuclear interactions. *Trends Genet* 23(2):81–90.
14. Werren JH, Baldo L, Clark ME (2008) *Wolbachia*: Master manipulators of invertebrate biology. *Nat Rev Microbiol* 6(10):741–751.
15. Cordaux R, Bouchon D, Grève P (2011) The impact of endosymbionts on the evolution of host sex-determination mechanisms. *Trends Genet* 27(8):332–341.
16. Rigaud T, Juchault P, Mocquard JP (1997) The evolution of sex determination in isopod crustaceans. *BioEssays* 19:409–416.
17. Bouchon D, Cordaux R, Grève P (2008) Feminizing *Wolbachia* and the evolution of sex determination in isopods. *Insect Symbiosis*, eds Bourtzis K, Miller T (Taylor and Francis, Boca Raton, FL), Vol 3, pp 273–294.
18. Martin G, Juchault P, Legrand JJ (1973) Mise en évidence d'un micro-organisme intracytoplasmique symbiote de l'oniscoïde *Armadillidium vulgare* Latr., dont la présence accompagne l'intersexualité ou la féminisation total des mâles génétiques de la lignée thélygène. *C R Acad Sci Paris* 276:2213–2216.
19. Cordaux R, Michel-Salzat A, Frelon-Raimond M, Rigaud T, Bouchon D (2004) Evidence for a new feminizing *Wolbachia* strain in the isopod *Armadillidium vulgare*: Evolutionary implications. *Heredity (Edinb)* 93(1):78–84.
20. Rigaud T (1997) Inherited microorganisms and sex determination of arthropod hosts. *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*, eds O'Neill SL, Hoffmann AA, Werren JH (Oxford Univ Press, New York), pp 81–101.
21. Legrand JJ, Juchault P, Heily G, Le Bote C (1984) Nouvelles données sur le déterminisme génétique et épigénétique de la monogénie chez le crustacé isopode terrestre *Armadillidium vulgare* Latr. *Genet Sel Evol* 16(1):57–84.
22. Juchault P, Rigaud T, Mocquard JP (1993) Evolution of sex determination and sex-ratio variability in wild populations of *Armadillidium vulgare* (Latr.) (crustacea, isopoda): A case study in conflict resolution. *Acta Oecologica* 14(4):547–562.
23. Cordaux R, Michel-Salzat A, Bouchon D (2001) *Wolbachia* infection in crustaceans: Novel hosts and potential routes for horizontal transmission. *J Evol Biol* 14:237–243.
24. Kondo N, Nikoh N, Ijichi N, Shimada M, Fukatsu T (2002) Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect. *Proc Natl Acad Sci USA* 99(22):14280–14285.
25. Dunning Hotopp JC, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317(5845):1753–1756.
26. Koutsovoulos G, Makepeace B, Tanya VN, Blaxter M (2014) Palaeosymbiosis revealed by genomic fossils of *Wolbachia* in a strongyloidean nematode. *PLoS Genet* 10(6):e1004397.
27. Thézé J, Leclercq S, Moumen B, Cordaux R, Gilbert C (2014) Remarkable diversity of endogenous viruses in a crustacean genome. *Genome Biol Evol* 6(8):2129–2140.
28. Pannunzio NR, Li S, Watanabe G, Lieber MR (2014) Non-homologous end joining often uses microhomology: Implications for alternative end joining. *DNA Repair (Amst)* 17:74–80.
29. Sinha S, Villarreal D, Shim EY, Lee SE (2016) Risky business: Microhomology-mediated end joining. *Mutat Res* 788:17–24.
30. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
31. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18(5):821–829.
32. Luo R, et al. (2012) SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1(1):18.
33. Gnerre S, et al. (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci USA* 108(4):1513–1518.
34. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27(4):578–579.
35. Boetzer M, Pirovano W (2012) Toward almost closed genomes with GapFiller. *Genome Biol* 13(6):R56.
36. Sommer DD, Delcher AL, Salzberg SL, Pop M (2007) Minimus: A fast, lightweight genome assembler. *BMC Bioinformatics* 8:64.
37. Kearse M, et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649.
38. Seemann T (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30(14):2068–2069.
39. Camacho C, et al. (2009) BLAST+: Architecture and applications. *BMC Bioinformatics* 10:421.
40. Lindner MS, Kollock M, Zickmann F, Renard BY (2013) Analyzing genome coverage profiles with applications to quality control in metagenomics. *Bioinformatics* 29(10):1260–1267.
41. Baldo L, et al. (2006) Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl Environ Microbiol* 72(11):7098–7110.
42. Paraskevopoulos C, Bordenstein SR, Wernegreen JJ, Werren JH, Bourtzis K (2006) Toward a *Wolbachia* multilocus sequence typing system: Discrimination of *Wolbachia* strains present in *Drosophila* species. *Curr Microbiol* 53(5):388–395.
43. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12):2725–2729.
44. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59(3):307–321.
45. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* 16(6):276–277.
46. Duplouy A, et al. (2013) Draft genome sequence of the male-killing *Wolbachia* strain wBol1 reveals recent horizontal gene transfers from diverse sources. *BMC Genomics* 14:20.
47. Krzywinski M, et al. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* 19(9):1639–1645.
48. Rigaud T, Juchault P (1993) Conflict between feminizing sex ratio distorters and an autosomal masculinizing gene in the terrestrial isopod *Armadillidium vulgare* Latr. *Genetics* 133(2):247–252.

EVOLUTION