Investigating the Molecular Genetic Basis of Cytoplasmic Sex Determination Caused by *Wolbachia* Endosymbionts in Terrestrial Isopods

Myriam Badawi [ID]**, Bouziane Moumen, Isabelle Giraud, Pierre Grève and Richard Cordaux** *

Laboratoire Ecologie et Biologie des Interactions, Equipe Ecologie Evolution Symbiose, Université de Poitiers, UMR CNRS 7267, Bât. B8, 5 rue Albert Turpin, TSA 51106, 86073 Poitiers CEDEX 9, France; Myriam.Badawi@univ-lemans.fr (M.B.); bouziane.moumen@univ-poitiers.fr (B.M.); isabelle.giraud@univ-poitiers.fr (I.G.); pierre.greve@univ-poitiers.fr (P.G.)

* Correspondence: richard.cordaux@univ-poitiers.fr; Tel.: +33-5-49-45-36-51

Abstract: In animals, sexual differences between males and females are usually determined by sex chromosomes. Alternatively, sex may also be determined by vertically transmitted intracellular microbial endosymbionts. The best known cytoplasmic sex manipulative endosymbiont is *Wolbachia* which can, for instance, feminize genetic males into phenotypic females in the terrestrial isopod *Armadillidium vulgare*. However, the molecular genetic basis of cytoplasmic sex determination is unknown. To identify candidate genes of feminization induced by *Wolbachia* strain *w*VulC from *A. vulgare*, we sequenced the genome of *Wolbachia* strain *w*Con from *Cylisticus convexus*, the most closely related known *Wolbachia* strain to *w*VulC that does not induce feminization, and compared it to the *w*VulC genome. Then, we performed gene expression profiling of the 216 resulting *w*VulC candidate genes throughout host developmental stages in *A. vulgare* and the heterologous host *C. convexus*. We identified a set of 35 feminization candidate genes showing differential expression during host sexual development. Interestingly, 27 of the 35 genes are present in the *f* element, which is a piece of a feminizing *Wolbachia* genome horizontally transferred into the nuclear genome of *A. vulgare* and involved in female sex determination. Assuming that the molecular genetic basis of feminization by *Wolbachia* and the *f* element is the same, the 27 genes are candidates for acting as master sex determination genes in *A. vulgare* females carrying the *f* element.

Keywords: cytoplasmic sex factor; feminization; sexual development; genome sequencing; gene expression profile; *Wolbachia*; terrestrial isopods

1. Introduction

Sex determination is a key biological pathway which governs the sexual differentiation of an individual into a male or female and its ability to produce male or female gametes. Although development as either a male or female is a generally conserved feature in animals, there is an amazing diversity of modes of sex determination [1,2]. The most common mode of sex determination is genotypic sex determination, in which nuclear genetic elements control sex determination. Prime examples of genotypic sex determination are sex chromosomes, i.e., chromosomes that carry sex-determining factors. Sex chromosome systems can take two major forms: male heterogamety (when males are heterozygous for the sex chromosomes and females homozygous, called XY systems), and female heterogamety (when females are heterozygous for the sex chromosomes and males homozygous, called ZW systems). The second mode of sex determination is environmental sex determination, in which external stimuli control sex determination. Environmental sex determination is defined here in

a very broad sense, meaning any factor that is not nuclear, be it abiotic or biotic, and can trigger male or female differentiation from a single nuclear genotype. Examples of non-nuclear triggers include temperature, photoperiod, social ties and parasites, among others.

Among parasites, several intracellular microorganisms have evolved the ability to feminize their hosts, a phenomenon known as cytoplasmic sex determination [3–5]. Disrupting the mode of sex determination of their hosts in favor of females may be advantageous for such inherited endosymbionts because they are predominantly transmitted vertically through female egg cytoplasm, not male sperm. Thus, males represent dead ends for these microorganisms. Consequently, any effect of the endosymbionts that distort host sex ratio towards females will be selectively advantageous for the endosymbionts, provided that the sex ratio distortion does not unduly affect the ability of hosts to reproduce. The best known causative agent of cytoplasmic sex determination is the α-proteobacterium *Wolbachia*, which has attracted considerable interest, notably for its wide range of hosts and induced phenotypes [6,7]. In particular, several *Wolbachia* strains have been shown to cause feminization of genetic males into phenotypic females [3,4,8]. Feminization has been reported in some insects [9,10] but is most widespread in terrestrial isopod crustaceans, including the common pillbug *Armadillidium vulgare* [11–13].

A. vulgare has a female heterogametic system of sex chromosomes (ZZ/ZW) [14]. In this context, inherited *Wolbachia* endosymbionts lead to the feminization of ZZ genetic males that develop into fully functional females during sexual differentiation [4,8,11,12]. As *Wolbachia* invades the population, it leads to the disappearance of heterogametic females (ZW) and therefore, to W sex chromosome extinction. Effectively, this means that the female sex-determining factor shifts from nuclear (W sex chromosome) to cytoplasmic (*Wolbachia*) localization [15]. Interestingly, *Wolbachia* horizontally transferred its genome into the host nucleus, in a genomic region associated with female sex-determination called the *f* element [16]. The *f* element ultimately led to the birth of a new W sex chromosome [16]. Thus, *Wolbachia* is not only a factor inducing cytoplasmic sex determination, it also has the potential to act as a substantial evolutionary force that can trigger turnovers of sex chromosomes in *A. vulgare*. However, the mechanisms and genes responsible for feminization by *Wolbachia* and the *f* element are unknown [15].

In general, the molecular mechanisms underlying *Wolbachia*-host interactions remain poorly known (reviewed in [17]). This is largely due to the fact that *Wolbachia* is an unculturable bacterium with an obligate intracellular lifestyle. With respect to feminization, genome sequencing of the feminizing *Wolbachia* strain *w*VulC identified many genes that could potentially interact with its host *A. vulgare*, especially genes that encode for proteins containing eukaryotic-like domains (at least 83 putative genes including 75 with ankyrin domain; Figure S1). These domains are commonly found in eukaryotic proteins and are relatively rare in bacteria, but particularly abundant in *Wolbachia* genomes [18]. Genes containing eukaryotic-like domains are particularly enriched in prophage WO, which is present in many *Wolbachia* genomes, sometimes in multiple copies [19–21]. These genes are of particular interest as their role in host-pathogen interactions has been demonstrated in other microorganisms such as *Legionella pneumophila* [22,23], *Coxiella burnetii* [22,24] and *Anaplasma phagocytophilum* [25]. In *Wolbachia*, it has already been shown that some of these genes show sex-specific expression [26] or are correlated to the induced phenotype [27–29]. In several *Wolbachia* strains that infect terrestrial isopods, the ankyrin gene *pk2b2* (embedded in one of the 7 prophage regions of the *w*VulC genome; Figure S1) is only expressed in females from species infected by feminizing strains [30]. Sequencing of the *w*VulC genome also revealed the presence of numerous export and secretion systems such as T4SS, T1SS, Sec, Tat, Lol and Bam, through which effectors may transit towards host cytoplasm. This result is consistent with earlier reports, indicating that the presence of such systems is a general feature of *Wolbachia* endosymbionts [31–33]. In *w*VulC, the T4SS has been well studied [31,34]; however, no effector or molecular mechanism of feminization induced by *Wolbachia* has been uncovered to date. This is in contrast with cytoplasmic incompatibility (CI), the most common phenotype induced by *Wolbachia*, which has recently been shown to involve the pair of genes, *cifA* and *cifB* [35–37].

So far, many studies aiming at identifying candidate genes involved in *Wolbachia*-host interactions used targeted approaches, i.e., by focusing on genes containing functional domains with prior knowledge of involvement in molecular interactions. Considering the quite limited amount of functional knowledge of *Wolbachia* genes (e.g., 23% of hypothetical genes in *w*VulC genome), it may be relevant to consider alternative strategies to identify candidate effectors to elucidate molecular mechanisms of induced phenotypes [38]. In this study, we used a combination of comparative genomics and expression profiles throughout host development to identify candidate genes of feminization induced by *Wolbachia w*VulC, without prior functional knowledge of genes. Specifically, we sequenced the genome of *Wolbachia w*Con from the terrestrial isopod *Cylisticus convexus* and compared it to the *w*VulC genome. *w*Con was selected because it is the most closely related known strain to *w*VulC available in our laboratory that does not induce feminization (but CI) [39,40]. In addition, it has been shown that *w*VulC also induces feminization when transfected into *C. convexus* and *w*Con also induces CI when transfected into *A. vulgare* [39,41], indicating that the observed phenotypes are related to the *Wolbachia* strains. Next, we performed gene expression profiling of selected *w*VulC genes throughout host developmental stages in *w*VulC native host *A. vulgare* and following *w*VulC transfection in a heterologous host (*C. convexus*). *C. convexus* was selected for comparison with *A. vulgare* because: (i) *w*VulC has been shown to induce feminization after transfection in *C. convexus*, and (ii) *A. vulgare* and *C. convexus* differ in their timing of sexual differentiation [41]. Using this strategy, we identified a set of 35 candidate genes out of the 1888 *w*VulC genes that may be implicated in feminization induced by *Wolbachia w*VulC in the host *A. vulgare*, two of which appear to be particularly promising.

2. Materials and Methods

2.1. Terrestrial Isopod Lines

All *C. convexus* individuals naturally infected with *Wolbachia* strain *w*Con were from our laboratory line CCw, derived from individuals caught in Avanton, France, in 2004. All *C. convexus* individuals carrying *Wolbachia* strain *w*VulC were from our laboratory line AW (derived from individuals caught in Villedaigne, France, in 1997) and were experimentally infected with *Wolbachia* strain *w*VulC (from our laboratory line ZN, derived from individuals caught in Celles sur Belle, France, 1991), as described in Badawi et al. [41]. All *A. vulgare* individuals naturally infected with *Wolbachia* strain *w*VulC were from our laboratory line ZN. Isopods were reared at 20 °C with food ad libitum (dead lime tree leaves and carrots) under a natural photoperiod, except those in cross-breeding and juveniles which were reared under a 18L:6D photoperiod.

2.2. wCon DNA Enrichment and Genome Sequencing

To identify the *C. convexus* tissue naturally possessing the highest *Wolbachia w*Con density, DNA was extracted from ovaries, the nervous chain and the hemolymph of four *C. convexus* females using a standard phenol/chloroform extraction [42]. Proportions of *w*Con, mitochondrial and nuclear DNA were estimated with quantitative PCR targeting single copy genes *wsp*, cytochrome oxidase I (COI) and androgenic hormone (AH), respectively. Reactions and cycle conditions were as described in Le Clec'h et al. [43] and in Table S1. Copy numbers were estimated using calibration range curves as described in Le Clec'h et al. [43]. We assessed the ratios of the base pair number from mitochondria (R_{mt}) and nucleus (R_{nuc}) relative to the *w*Con genome, as follows: $R_{mt} = (G_{mt} \times N_{COI})/(G_{Wo} \times N_{wsp})$ and $R_{nuc} = (G_{nuc} \times N_{AH})/(G_{Wo} \times N_{wsp})$, where G_{mt}, G_{nuc} and G_{Wo} are mitochondrial, nuclear and *Wolbachia* genome sizes, respectively, and N_{COI}, N_{AH} and N_{wsp} are COI, AH and *wsp* copy numbers, respectively. Absolute genome sizes were estimated as 1.7 Mb and 2 Gb for *w*Con and nuclear genomes respectively, based on information available for the closely related *Wolbachia w*VulC/ *A. vulgare* complex (http://www.genomesize.com/). *C. convexus* mitochondrial genome length is 14 kb, based on sequencing data [44].

Ovaries were selected for further analyses as they contained the highest ratio of *w*Con copies. We dissected ovaries, as well as nerve cords of 30 *C. convexus* sisters infected with the *w*Con *Wolbachia* strain. First, we performed PCR amplification and Sanger sequencing of the *Wolbachia*-specific markers *wsp* and *ftsZ* using the DNA of the nerve cords from each of the 30 females as templates and then verified that the *Wolbachia* strain was indeed *w*Con (see Table S1 for primers and PCR conditions). Next, to enrich the sample in *w*Con DNA, we homogenized the ovaries of the 30 females with a Dounce tissue grinder B within a PBS solution containing sucrose (0.25 M) and L-glutamine (5 mM) to crush cells while keeping nuclei intact. We then passed the solution through a 5 μm filter that specifically retained nuclei but not *Wolbachia*. Next, the DNA of the *Wolbachia*-enriched solution was extracted using a standard phenol/chloroform procedure. RNA contaminants were removed with RNase A treatment (0.2 μg/μL at 37 °C for 1 h). Then, the sample was purified through another round of phenol/chloroform extraction. We determined the relative proportion of *w*Con DNA relative to mitochondrial and nuclear DNA with quantitative PCR (qPCR), as described above. The DNA sample was used by GenoScreen (Lille, France) to prepare a Nextera sequencing library. The library was sequenced by GenoScreen in a 1/4 and a 1/8 454 GS FLX sequencer runs with Titanium chemistry.

Bad quality reads and mitochondrial reads were filtered out and the remaining reads were assembled de novo with gsAssembler (Newbler 2.6) with default parameters except seed step set to 1, yielding 5153 contigs. *w*Con contigs were retrieved using BLAT (minimal identity set to 60), Mauve and R2CAT [45–47]. Six representative *Wolbachia* reference genomes from supergroups A and B were used to retrieve *w*Con contigs: *w*Mel (NC_002978.6), *w*Ri (NC_012416.1), *w*Pip-Pel (NC_010981.1), *w*AlB (NZ_CAGB00000000.1), *w*VitB (NZ_AERW00000000.1) and *w*VulC (ALWU00000000). The assembly of 389 contigs showing high similarity with *Wolbachia* genomes were manually improved by exploiting the de Bruijn graph to join contigs truncated due to repeats, which resulted in a final *w*Con assembly of 237 contigs. Gene prediction and annotation was performed with PROKKA pipeline [48] and manually curated in artemis. Clusters of Orthologous Groups of proteins (COG) were assigned to each predicted protein using BLASTp with minimal identity of 70% on at least 70% of the protein sequences. The annotated genome sequence and raw sequence data were deposited in GenBank under bioproject accession number PRJNA439208.

2.3. In Silico Classification of wVulC Genes

To determine a list of feminization candidate genes, all annotated genes from the *w*VulC genome (GenBank accession number ALWU00000000) were processed using a homemade pipeline of Perl scripts (Supplementary Materials 1). First, according to genome annotation, all repeats (including transposable elements, genes coding for prophages and other repeated genes), pseudogenes (including all coding sequences of genes annotated as "pseudogenes" and small coding sequences of genes split in two chunks) and small peptides (<50 amino acids) were removed. Repeats were discarded because they were not well suited for downstream expression analyses by PCR. Pseudogenes and small peptides were discarded as they are unlikely to be functional genes. Second, all genes belonging to the core genome of *Wolbachia* strains that infect arthropods were discarded. This core genome was computed using OrthoMCL [49] with a minimal identity of 70% and a minimal e-value of 10^{-6}, using the following five complete *Wolbachia* genomes from supergroups A and B: *w*Mel (NC_002978.6), *w*Ri (NC_012416.1), *w*Ha (NC_021089.1), *w*No (NC_021084.1), *w*Pip-Pel (NC_010981.1). Finally, all proteins from *w*VulC showing 100% similarity with their homolog in the closely related, non-feminizing *w*Con genome on at least 90% of their length, based on BLASTp searches, were discarded. The remaining genes were considered as candidates and were selected for expression analyses.

2.4. Expression of wVulC Candidate Genes throughout Host Sexual Development

To investigate the expression of *Wolbachia* candidate genes during host sexual development, genetic crosses involving *A. vulgare* and *C. convexus* individuals were performed between uninfected males and females carrying the *w*VulC *Wolbachia* strain. In the terrestrial isopods *A. vulgare* and

C. convexus, embryos that inherited the feminizing *Wolbachia* strain *w*VulC develop into functional females. This conversion of genetic males into functional phenotypic females occurs during host sexual differentiation. In brief, *A. vulgare* juvenile development occurs within a period of 10 to 15 weeks after the release of juveniles from the female ventral pouch. During this period, eight post-embryonic stages were defined, each corresponding to an intermolt stage. Gonads differentiate during stages 4 to 6 and after this period the experimental reversion of gonadal sex becomes impossible [50–52]. Therefore, *Wolbachia* is assumed to act before or during sexual differentiation as it inhibits male gonad differentiation and hence converts genetic males into phenotypic females. *C. convexus* has a different sexual differentiation timing compared to *A. vulgare*, as it shows a one-stage shift and occurs during stages 3 to 5 [41]. After hatching, juveniles were sampled individually 2–3 days after each molt, from the first molt (stage 1) to the seventh molt (stage 7), and stored in liquid nitrogen.

First, we tested whether *w*VulC candidate genes are expressed in any of the gonadal differentiation key stages (3 to 6) in *A. vulgare* juveniles. We co-extracted DNA and RNA of 3, 2, 2 and 2 juvenile pools sampled from stages 3 to 6 respectively, using the AllPrep DNA/RNA kit (Qiagen®, Venlo, The Netherlands), according to the manufacturer's instructions. We confirmed *w*VulC presence in each sample by amplifying molecular markers *gatB* and *wsp* by PCR (Table S1 for primers and PCR conditions). RNA samples were then treated with DNase I (4u at 37 °C for 1h and inactivated at 75 °C for 10 min, New England Biolabs®, Ipswich, MA, USA) and cDNA synthesis was performed using the Superscript SSIII kit (Invitrogen®, Carlsbad, CA, USA), according to the manufacturer's instructions, and using 250 ng of RNA (estimated with Nanodrop, Thermo Fisher, Waltham, MA, USA). Then, RT-PCR was performed on all candidate genes (Table S2 for primers and PCR conditions) and PCR products were run on an agarose gel 1.5%. After staining the gel for 10 min in an ethidium bromide bath, bands were revealed under UV light. Candidate genes being expressed during at least one host developmental stage were selected for expression quantification.

*Wolbachia w*VulC gene expression was quantified by quantitative real-time PCR (qRT-PCR) on three successive samplings. The first sampling was composed of 40, 24, 16, 16 and 16 *A. vulgare* juvenile pools sampled from stages 2 to 6, respectively. The second sampling was composed of triplicated pools of 10, 10, 9, 6, 6, 6 and 6 *A. vulgare* juveniles corresponding to stages 1 to 7, respectively. The third sampling was composed of triplicated pools of eight *C. convexus* juveniles for each stage from 1 to 7. All quantitative experiments were conducted with two technical replicates, using a LightCycler LC480 device (Roche®, Basel, Switzerland). Reactions were made in a total volume of 10 μL composed of 1× Fast SYBR GREEN Master Mix (Roche®, Basel, Switzerland), 5 μM of each primer, and 2.5 μL of 4×-diluted cDNA synthesized as above. Cycle conditions were as described in Le Clec'h et al. [43].

To identify appropriate reference genes, we measured the expression of 15 *Wolbachia* housekeeping genes (*wsp*, *coxA*, *atpD*, *sucB*, *gatB*, *gltA*, *L2*, *L20*, *S4*, *fabF*, *pyrB*, *purF*, *tkt*, *hcpA* and *fbpA*) by qRT-PCR in *A. vulgare* stages 2 to 6. We then selected those exhibiting a highly correlated expression ($r > 0.95$) using the BestKeeper software (version 1) [53] (Table S1 for primers and PCR conditions). At last, we verified that the expression of the selected genes remained constant across stages in *A. vulgare* and *C. convexus* relative to the *wsp* copy number. This procedure resulted in the validation of six genes: *gltA*, *L2*, *L20*, *S4*, *fabF* and *purF*. Expression ratios were calculated by combining efficiency and $2^{\Delta\Delta ct}$, according to LightCycler 480 SW v1.5 software manual (Roche®, Basel, Switzerland). Efficiency and cycle threshold were assessed by real-time PCR (RT-PCR) Miner [54]. Then, each expression ratio was calibrated against stage 2 (corresponding to the earliest stage for which all genes were amplified in all samplings). Under- and over-expression were considered with thresholds of 0.5 and 2, respectively for all successive samplings.

To compare *w*VulC gene expression patterns between *A. vulgare* and *C. convexus* hosts, we used cross-correlation tests using R software [55]. This test measures correlation by taking into account gene expression patterns shifted between stages (hence taking into account the one-stage shift of sexual differentiation that occurs between the two hosts [41]). This test was applied after an autocorrelation test aimed at verifying the absence of autocorrelation between the different time steps. In positive cross-correlation tests, we computed the η^2 value to verify that the variance did not influence the cross-correlation test.

2.5. Functional Annotation of Candidate Genes

Function prediction of the final set of candidate genes was performed by searching the following databases: PFAM [56], SMART [57], cdd [58], cog [59], kog [60] and signal peptide [61]. Results were accounted as significant with an e-value $<10^{-4}$.

2.6. Identification of Candidate Genes in the f Element

Considering the *f* element as the *Wolbachia* nuclear insert characterized by Leclercq et al. [16], the presence of the final set of feminization candidate genes from *w*VulC in the *f* element was investigated using BLASTp against *f* element sequences sensu Leclercq et al. [16] (GenBank accession numbers LYUU01002088.1-LYUU01002096.1), with a maximal e-value set to 0.001 and at least 90% of identity.

3. Results

3.1. Sequencing of wCon Wolbachia Genome

To sequence the *w*Con *Wolbachia* genome, we first applied an enrichment method, as *Wolbachia* DNA is often in low abundance compared to host genetic material (nuclear and mitochondrial). To evaluate which *C. convexus* tissue contains the highest abundance in *Wolbachia* DNA, we quantified three single-copy genes by means of qPCR (androgenic hormone for nuclear DNA, cytochrome oxidase I for mitochondrial DNA and *wsp* for *Wolbachia* DNA) in three different tissues (gonads, hemolymph and nervous cord). The highest level of *w*Con DNA was found in gonads (~1 *Wolbachia* bp for ~330 nuclear bp), followed by the nervous cord (~1 *Wolbachia* bp for ~720 nuclear bp) and hemolymph (~1 *Wolbachia* bp for ~34,720 nuclear bp) (Figure 1A). These figures correspond to 1–4 *Wolbachia* cells per host cell on average. The quantity of mitochondrial nucleotides was lower than that of *Wolbachia* (~1 *Wolbachia* bp for ~0.4 mitochondrial bp in ovaries; Figure 1B). Subsequently, we applied our DNA extraction and enrichment protocol to a pool of 30 gonad pairs from *C. convexus* sisters, reaching a >300-fold enrichment in *w*Con DNA (Figure 1C,D).
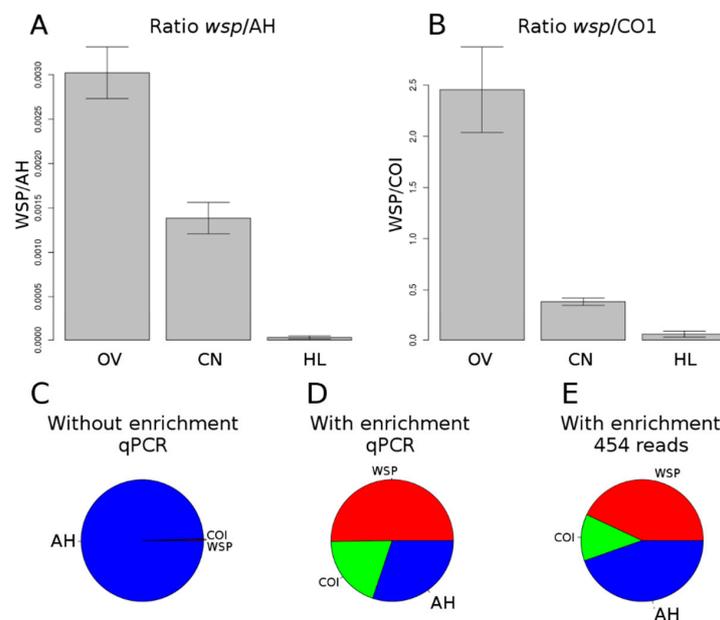

Figure 1. Ratios of *wsp* gene copy number relative to androgenic hormone (AH) copy number (**A**) and cytochrome oxidase I (COI) copy number (**B**), in ovaries (OV), nervous cord (CN) and hemolymph (HL). Proportions of base pairs based of *w*Con (red), mitochondrial (green) and nuclear genomes (blue) estimated by quantitative PCR (qPCR) without (**C**) and with (**D**) enrichment, and by mapping of pyrosequencing reads on assembled genomes (**E**). WSP: please define.

Sequencing of this sample generated 395,850 reads representing a total of 110,934,433 bp. 16,491 reads were filtered out for low quality and 47,065 reads that mapped against *C. convexus* mitochondrial genome were discarded. In total, 163,099 reads (43%) were mapped against the final set of *w*Con contigs. Therefore, qPCR estimation of *w*Con DNA proportion (50%) was close to that obtained by sequencing (Figure 1D,E). From these reads, we assembled the *w*Con *Wolbachia* genome in 237 contigs, with a total size of 2.1 Mb, an average sequencing depth of 25×, a N50 of 4.5 kb and a GC content of 34.7%. Contig size ranged from 404 to 71,427 bp. PROKKA annotation pipeline revealed the presence of 2359 coding sequences with an average size of 702 bp per gene, 35 transfer RNA (tRNA) and 3 ribosomal RNA (rRNA) genes (Table 1). Genome size is likely overestimated since redundant repeats are present at the ends of contigs resulting from manual scaffolding, as described in Section 2.2. Consistently, assignment of annotated coding DNA sequences (CDS) to COG categories revealed that *w*Con contains a high number of genes in the category "Replication, Recombination and DNA Repair" among the 7 *Wolbachia* genomes compared (Figure S2). This COG category includes transposable elements (402 repeats in *w*Con).

Core genome analysis performed on five complete *Wolbachia* genomes from arthropods (*w*Mel, *w*Pip-Pel, *w*Ri, *w*Ha and *w*No) revealed the presence of 762 orthologous groups in their core genome. We found 752 and 755 of these orthologous groups in the *w*Con and *w*VulC genomes, respectively (748 of which are shared by *w*Con and *w*VulC). Furthermore, 35 tRNA and 3 rRNA genes were characterized in both *w*Con and *w*VulC genomes (Table 1). Altogether, these analyses indicate that we obtained reliable assemblies of the *w*Con and *w*VulC genomes, capturing most of the genetic information despite assembly fragmentation (237 and 10 contigs for *w*Con and *w*VulC, respectively).

3.2. In Silico Determination of Candidate Genes

The *w*VulC *Wolbachia* genome contains 1888 annotated coding sequences. A succession of in silico filters (see Materials and Methods) allowed us to establish a set of feminization candidate genes. As a result, based on the *w*VulC annotation, we excluded all repeats ($n = 792$), pseudogenes ($n = 26$), genes <150 bp ($n = 52$) and small CDS of genes split in two CDS ($n = 16$). We also removed genes orthologous to the ones identified in the core genome of five *Wolbachia* strains of arthropods (*w*Mel, *w*Ri, *w*Pip-Pel, *w*Ha and *w*No) ($n = 721$). This number is different from the above core genome analysis ($n = 755$) because a subset of core genes was discarded as part of previous filters. Next, remaining *w*VulC proteins were compared to their homologs in the non-feminizing *w*Con *Wolbachia* genome (CI-inducing strain), and those showing 100% similarity ($n = 16$) were excluded (Table S3, Figure S1). The 265 remaining *w*VulC genes were considered as a first set of feminization candidate genes and were processed for expression studies.

Table 1. *Wolbachia* genome features with *w*Con and *w*VulC.

Supergroup	B	B	B	B	A	A	A	C	D	F
Strain	*w*Con	*w*VulC	*w*No	*w*Pip-Pel	*w*Mel	*w*Ri	*w*Ha	*w*Oo	*w*Bm	*w*Cle
Host	*Cylisticus convexus*	*Armadillidium vulgare*	*Drosophila simulans*	*Culex quinque-fasciatus*	*Drosophila melanogaster*	*Drosophila simulans*	*Drosophila simulans*	*Onchocerca ochengi*	*Brugia malayi*	*Cimex lectularius*
Size (bp)	2,109,518	1,663,741	1,301,823	1,482,455	1,267,782	1,445,873	1,295,804	957,990	1,080,084	1,250,060
G + C%	34.69	34.49	34.50	34.60	35.46	35.40	35.34	32.10	35.20	36.30
Coding sequences	2356	1888	1040	1275	1195	1150	1010	842	805	1216
Coding base density	0.78	0.83	0.80	0.86	0.94	0.80	0.78	0.67	0.75	0.75
Gene average size	702	725	1013	951	851	976	1000	757	899	771
rRNA	3	3	3	3	3	3	3	3	3	3
tRNA	35	35	34	34	34	34	34	34	34	34
Pseudogenes	NA	375	95	110	74	114	93	196	98	109

3.3. Candidate Gene Expression throughout Host Development

Among the 265 remaining *w*VulC genes, successful PCR amplification was obtained for 216 candidate genes. Then, gene expression of the 216 candidate genes from the *w*VulC *Wolbachia* genome was investigated by RT-PCR throughout *A. vulgare* development, in particular during the stages of sexual differentiation, i.e., from stages 3 to 6. We found that 139 out of 216 *w*VulC genes were expressed during at least at one of these host developmental stages (Table S3). Next, gene expression of the 139 candidate genes was investigated by qRT-PCR throughout developmental stages of the host *A. vulgare,* using two successive samplings. First, we investigated expression profiles of these genes during host developmental stages 2 to 6, using stage 2 for calibration (see Materials and Methods, Figure S3). This analysis identified 13 and 29 genes that were 2-fold under and over-expressed, respectively, in at least one of stages 3 to 6. Expression profiles of these 42 differentially expressed genes were then investigated within the frame of a second sampling. The second sampling was designed to confirm and refine expression profiles of *w*VulC candidate genes throughout the full development sequence of the host *A. vulgare* (i.e., from stages 1 to 8). In total, seven genes were below the 2-fold expression threshold for all replicates. Therefore, 35 genes were consistently under or over-expressed, making them the final set of candidate genes.

Finally, expression profiles of the 35 *w*VulC candidate genes were investigated, during developmental stages 1 to 7 in the heterologous host *C. convexus*, in which *w*VulC transfection was shown to be associated with feminization (Figure 2) [41]. We found that all 35 *w*VulC genes were expressed during *C. convexus* development, but only 29 of these genes were significantly down or up-regulated in both hosts (Figure 2, Table S3). A cross-correlation analysis of gene expression profiles between the two hosts indicated that three *w*VulC genes exhibited the same gene expression profile throughout the entire developmental sequence in both hosts: wVul_1408 and wVul_1821 with one-stage shift that perfectly matches the one-stage shift of sexual differentiation existing between *A. vulgare* and *C. convexus* (Figure 2) [41], and wVul_0881, but without any stage shift between the two hosts (Figure S4).
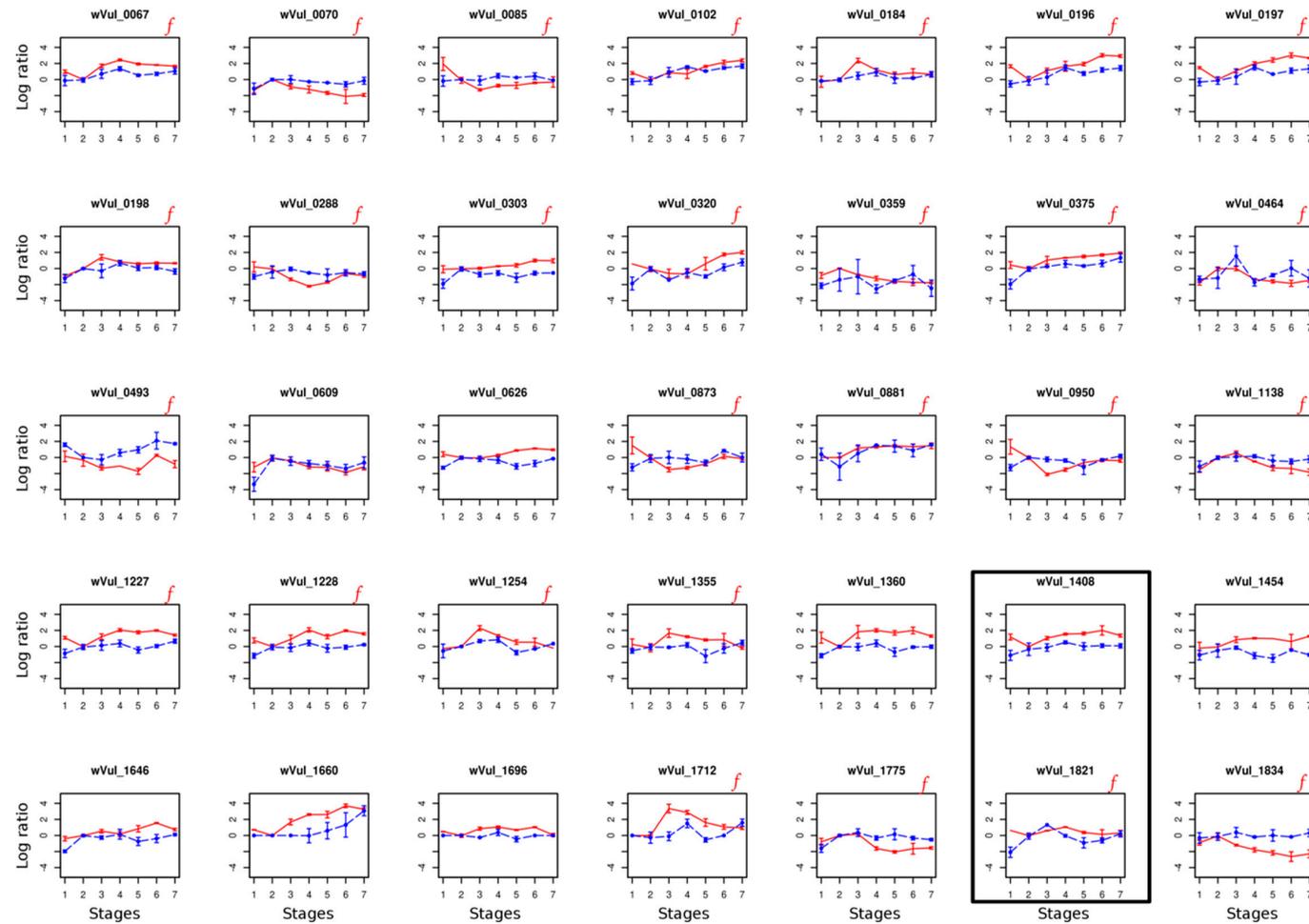
Figure 2. Candidate gene log2 expression of 35 candidate genes during both homologous and heterologous host development (*A. vulgare* in red and *C. convexus* in blue, respectively). All candidate genes are consistently over- or under-expressed during native host development (in *A. vulgare*). All genes present in the *f* element are labeled with a red "*f*". The two genes showing a significant cross-correlation with exactly one-stage shift between the two hosts are boxed.

3.4. Prediction of General Function of wVulC Candidate Genes

We investigated the general function of the 35 *w*VulC candidate genes that are consistently over or under-expressed throughout *A. vulgare* development. We were able to infer functions for 12 of these genes, in which we found a total of 14 different domains (up to 4 domains per gene) (Table 2; Table S4). We unraveled the following predicted domains associated with:

1.　Eukaryotic-like domains involved in protein-protein interactions: ANK (Ankyrin Repeat) [62,63] in wVul_0303, wVul_1408 and wVul_1775; LRR (Leucine Rich Repeat) [64,65] in wVul_0493; TPR (Tetratricopeptide Repeat) [66] in wVul_0881;

2.　Cytoskeleton: Hook in wVul_0067 that mediates attachment to microtubules [67]; Spc7 in wVul_0375 that is required for kinetochore-spindle association during cell division [68];

3.　DNA binding: Smc (Structural Maintenance of Chromosomes) in wVul_0067, wVul_0085 and wVul_0375 that binds DNA and acts in organization and segregation of chromosomes for partition [69,70]; AAA (ATPases Associated with diverse cellular Activities) [71] in wVul_0375 that are involved in diverse cellular functions including DNA replication, gene expression and also vesicle-mediated transport and peroxisome assembly; p53 interaction protein in wVul_1775 that is responsible for DNA binding [72].

4.　Hydrolase function: Lipase class 3 in wVul_0873 that hydrolyses ester linkages of triglycerides [73,74]; Peptidases/proteases that hydrolysepeptidic bounds [75,76] in wVul_1408, wVul_1454 and wVul_1775;

5.　Membranes: Mitofilin in wVul_0085 that is an inner membrane protein domain of mitochondria that controls cristae morphology [77,78]; IncA in wVul_1360 that is associated with homotypic fusion of inclusions in the obligate intracellular bacterium *Chlamydia trachomatis* [79]; OmpA-like superfamily in wVul_0359 that is an outer membran protein domain [80]; Trp in wVul_1408 and wVul_1775 that is a transient-receptor-potential calcium channel protein [81].

Table 2. Functional domains found in the 35 candidate genes. Underlined genes are present in the *f* element.

Gene Number	Annotation	Brief Description
67; 85; 375	Smc	Chromosome Segregation
375	AAA	DNA breaks
375	Spc7	Chromosome Segregation
67	Hook	N-term binding to microtubules
85	Mitofilin	Mitochondria inner membrane protein
359	OMPA-like	Surface
1360	IncA	Recruitment of lipids
1408; 1775	Trp	Transient receptor channels
493	LRR	Eukaryotic-like domain
881	TPR	Eukaryotic-like domain
303; 1408; 1775	ANK	Eukaryotic-like domain
873	Lipase	Hydrolizetriglycerides
1408; 1454; 1775	Peptidase/protease	Hydrolise peptide bonds
1775	p53	Bind to DNA, eukaryotic-like domain

3.5. Presence of Candidate Genes in the f Element

We searched the *f* element sequence for the 35 *w*VulC candidate genes. Blast results indicated that 27 genes are present in the *f* element with copy numbers varying from 1 to 5 (Table S5). The 8 *w*VulC candidate genes with no hit in the *f* element were wVul_0609, wVul_0626, wVul_1360, wVul_1408, wVul_1454, wVul_1646, wVul_1660 and wVul_1696. Among the 27 genes present in both *w*VulC and the *f* element, 20 presented identical copies at the amino acid level (Table S5). The lowest amino acid identity was observed for wVul_0198 (95.35% identity). There was no significant enrichment of differentially expressed genes in the *f* element (Fisher's Exact test, $p = 0.22$).

4. Discussion

Our analyses resulted in the identification of a reduced set of 35 candidate genes that may be implicated in *Wolbachia w*VulC-mediated feminization in the native host *A. vulgare* and the heterologous host *C. convexus*. The vast majority of *w*VulC genes that we tested are consistently expressed throughout the entire host development sequence, as previously observed for *Wolbachia w*Mel in its host *Drosophila melanogaster* [82]. Only one-fourth of genes analyzed by qRT-PCR (35/139) are expressed differentially, of which only one-third are functionally annotated (12/35). For example, we noted the presence of several eukaryotic-like domain proteins (ANK, TPR and LRR) and an outer membrane protein (OmpA-like) in our set of candidate genes. Identifying genes encoding such domains among our candidates is not really surprising, as previous host/*Wolbachia* molecular interaction studies have often searched for such functional domains when using targeted strategies to identify candidate genes [27–29]. One advantage of the strategy we used here was the ability to identify candidate genes without any a priori knowledge of their function. This allowed us to identify functionally annotated candidate genes among the "usual suspects", as well as candidate genes for which no functional information is available. Genes from the latter category may be prime choices for future functional studies, in the context of feminization induced by *w*VulC. This point may be relevant in identifying *Wolbachia* effectors, as exemplified by *cifA* and *cifB*, and their homologs *cidA* and *cidB*, which were not functionally annotated before their identification as CI-effectors in *w*Mel and *w*Pip-Pel strains, respectively [35,36].

The candidate gene approach we used also has its own potential limitations. First, as it largely relies on gene expression patterns, we implicitly assumed that gene expression and protein expression are correlated, which may not necessarily be true. This assumption is, however, generally made by studies involving gene expression. Second, we opted to discard core genes and mobile genetic elements. In particular, the rationale for removing mobile genetic elements was their repeated nature that makes them poorly suited to the design of specific PCR assays. Nevertheless, it is clear that genes encoded by mobile genetic elements may be involved in host-*Wolbachia* interactions, as perfectly exemplified by the *cifA* and *cifB* genes that are located in prophage WO [35,36]. However, our strategy did not dismiss those genes as candidate genes. Instead, we opted to give them lower priority in our expression analyses. Third, our approach is based on gene expression profiling throughout host development. However, during host development, *Wolbachia* is likely involved in various biological processes unrelated to feminization. For example, while host cells are dividing, *Wolbachia* cell population is also in expansion, replicating fast and colonizing new cells [83–88]. Therefore, a subset of our candidate genes may have passed our successive filters not because they are involved in feminization but because they are involved in a confounding biological process taking place during host sexual development.

In this context, the three domains Smc, Hook and Spc7 found in the wVul_0067, wVul_0085 and wVul_0375 candidate genes, are consistent with the involvement of these candidate genes in cell division [67–70]. Moreover, *Wolbachia* cell multiplication requires lipid recruitment to generate new membranes and additional nutrient supply. Thus, it is conceivable that the wVul_1360 candidate gene containing an IncA domain may contribute to lipid recruitment, as it seems to act in *C. trachomatis* [79,89,90]. This may also explain the presence of several lipase or peptidase domains among our candidate genes, which may contribute to additional nutrient supply. Interestingly, two candidate genes (wVul_1408 and wVul_1775) possess both a eukaryotic-like domain (ANK) and a peptidase/protease domain, suggesting that *Wolbachia* might cleave host peptides.

Among the 35 feminization candidate genes we identified, 27 were also found in the *f* element, which is a piece of a feminizing *Wolbachia* genome horizontally transferred in the nuclear genome of *A. vulgare* and involved in female sex determination [16]. Assuming that the molecular genetic basis of feminization by *Wolbachia* and the *f* element is the same, the 27 aforementioned genes are candidates for acting as master sex determination genes in the *A. vulgare* genome. However, in spite of the high nucleotide similarity of these genes in *w*VulC and the *f* element, their function or regulation may not necessarily be conserved in the two genetic entities. This may be the case especially when considering

that *Wolbachia* genes are located in the cytoplasm and controlled in a prokaryotic context while *f* element genes are located in the nucleus and controlled in a eukaryotic context. Furthermore, feminization induced by *Wolbachia* and the *f* element presents differences (reviewed in [15]). For example, it is possible to experimentally masculinize young females carrying the *f* element by grafting of androgenic glands [91]. By contrast, young females infected by *Wolbachia* are not masculinized using the same experimental procedure [92,93]. Thus, the underlying mechanisms of feminization by the *f* element and *Wolbachia* may differ, at least partly. Nevertheless, given that the *f* element is derived from *Wolbachia*, it is parsimonious to consider that the molecular genetic basis for feminization shows at least some overlap between the two feminizers. For this reason, it is relevant to consider the 27 *Wolbachia* feminization candidate genes present in the *f* element as candidate genes for master sex determination in females carrying the *f* element. In this context, it will be useful to investigate the expression profiles of these candidate genes during sexual development of *A. vulgare* females carrying the *f* element.

In conclusion, we identified a set of 35 *Wolbachia* *w*VulC genes that may be involved in the induction of cytoplasmic sex determination in the terrestrial isopod *A. vulgare*. Our results open the avenue for further functional analyses of these candidate genes, which may provide additional insights into the molecular genetic basis and cellular mechanisms of feminization. Of the 35 genes, 2 genes (*w*Vul_1408 and *w*Vul_1821) presented an expression profile in the heterologous host *C. convexus* which is correlated with that in the natural host *A. vulgare* and showing a one-stage shift corresponding to the one-stage shift existing in the sexual development sequence of *A. vulgare* and *C. convexus*. While *w*Vul_1821 contains no known functional domain, *w*Vul_1408 contains two known functional domains, including an Ankyrin Repeat Domain and a protease domain. Although *w*Vul_1408 is absent from the *f* element, it nevertheless constitutes a prime candidate for being involved in *Wolbachia* *w*VulC-mediated feminization. While *w*Vul_1408 may constitute a top candidate in subsequent studies aiming at unraveling the molecular genetic basis of feminization, it may also be important to invest efforts in the identification of the function of other candidate genes such as *w*Vul_1821. This will ultimately allow us to better understand the impact of *Wolbachia* endosymbionts on the mechanisms of sex determination of their terrestrial isopod hosts. It will also shed new light on the role of *Wolbachia* in sex chromosome turnovers experienced by *A. vulgare* in particular [16], and terrestrial isopods in general [94].

References

1. Bachtrog, D.; Mank, J.E.; Peichel, C.L.; Kirkpatrick, M.; Otto, S.P.; Ashman, T.-L.; Hahn, M.W.; Kitano, J.; Mayrose, I.; Ming, R.; et al. Sex determination: Why so many ways of doing it? *PLoS Biol.* **2014**, *12*, e1001899. [CrossRef] [PubMed]

2. Schwanz, L.E. The evolution of sex determination by Leo W. Beukeboom and Nicolas Perrin. *Q. Rev. Biol.* **2015**, *90*, 333–334. [CrossRef]

3. Cordaux, R.; Bouchon, D.; Grève, P. The impact of endosymbionts on the evolution of host sex-determination mechanisms. *Trends Genet.* **2011**, *27*, 332–341. [CrossRef] [PubMed]

4. Rigaud, T. Inherited microorganisms and sex determination of arthropod hosts. In *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*; O'Neill, S.L., Hoffmann, A.A., Werren, J.H., Eds.; Oxford University Press: Oxford, UK, 1997; pp. 81–101.

5. Jahnke, M.; Smith, J.E.; Dubuffet, A.; Dunn, A.M. Effects of feminizing microsporidia on the masculinizing function of the androgenic gland in *Gammarus duebeni*. *J. Invertebr. Pathol.* **2013**, *112*, 146–151. [CrossRef] [PubMed]

6. Werren, J.H.; Baldo, L.; Clark, M.E. *Wolbachia*: Master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **2008**, *6*, 741–751. [CrossRef] [PubMed]

7. Zug, R.; Hammerstein, P. Still a host of hosts for *Wolbachia*: Analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PLoS ONE* **2012**, *7*, e38544. [CrossRef] [PubMed]

8. Bouchon, D.; Cordaux, R.; Grève, P. Feminizing *Wolbachia* and the evolution of sex determination in isopods. In *Insect Symbiosis*; Bourtzis, K., Miller, T., Eds.; Taylor and Francis Group LLC: Boca Raton, FL, USA, 2008; Volume 3, pp. 273–294.

9. Hiroki, M.; Kato, Y.; Kamito, T.; Miura, K. Feminization of genetic males by a symbiotic bacterium in a butterfly, *Eurema hecabe* (Lepidoptera: Pieridae). *Naturwissenschaften* **2002**, *89*, 167–170. [CrossRef] [PubMed]

10. Negri, I.; Pellecchia, M.; Mazzoglio, P.J.; Patetta, A.; Alma, A. Feminizing *Wolbachia* in *Zyginidia pullula* (Insecta, Hemiptera), a leafhopper with an XX/X0 sex-determination system. *Proc. Biol. Sci.* **2006**, *273*, 2409–2416. [CrossRef] [PubMed]

11. Cordaux, R.; Michel-Salzat, A.; Frelon-Raimond, M.; Rigaud, T.; Bouchon, D. Evidence for a new feminizing *Wolbachia* strain in the isopod *Armadillidium vulgare*: Evolutionary implications. *Heredity* **2004**, *93*, 78–84. [CrossRef] [PubMed]

12. Martin, G.; Juchault, P.; Legrand, J.-J. Mise en évidence d'un micro-organisme intracytoplasmique symbiote de l'oniscoïde *Armadillidium vulgare* latr. dont la présence accompagne l'intersexualité ou la féminisation totale des mâles génétiques de la lignée thélygène. *CR Acad Sci Paris Ser. D* **1973**, *276*, 2313–2316. (In French)

13. Rigaud, T.; Juchault, P.; Mocquard, J.P. Experimental study of temperature effects on the sex ratio of broods in Terrestrial Crustacea *Armadillidium vulgare* Latr. Possible implications in natural populations. *J. Evol. Biol.* **1991**, *4*, 603–617. [CrossRef]

14. Juchault, P.; Legrand, J.-J. Croisement de néo-mâles experimentaux chez *Armadillidium vulgare* Latr. (Crustace, Isopode, Oniscoide). Mise en evidence d'une hétérogamétie femelle. *CR Acad Sci Paris Ser. D* **1972**, *274*, 1387–1389. (In French)

15. Cordaux, R.; Gilbert, C. Evolutionary significance of *Wolbachia*-to-animal horizontal gene transfer: Female sex determination and the f element in the isopod *Armadillidium vulgare*. *Genes* **2017**, *8*, 186. [CrossRef] [PubMed]

16. Leclercq, S.; Thézé, J.; Chebbi, M.A.; Giraud, I.; Moumen, B.; Ernenwein, L.; Grève, P.; Gilbert, C.; Cordaux, R. Birth of a W sex chromosome by horizontal transfer of *Wolbachia* bacterial symbiont genome. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 15036–15041. [CrossRef] [PubMed]

17. Bhattacharya, T.; Newton, I.L. Mi Casa es Su Casa: How an intracellular symbiont manipulates host biology. *Environ. Microbiol.* **2017**. [CrossRef] [PubMed]

18. Al-Khodor, S.; Al-Quadan, T.; Abu Kwaik, Y. Temporal and differential regulation of expression of the eukaryotic-like ankyrin effectors of *Legionella pneumophila*. *Environ. Microbiol. Rep.* **2010**, *2*, 677–684. [CrossRef] [PubMed]

19. Gavotte, L.; Henri, H.; Stouthamer, R.; Charif, D.; Charlat, S.; Boulétreau, M.; Vavre, F. A Survey of the bacteriophage WO in the endosymbiotic bacteria *Wolbachia*. *Mol. Biol. Evol.* **2007**, *24*, 427–435. [CrossRef] [PubMed]

20. Kent, B.N.; Bordenstein, S.R. Phage WO of *Wolbachia*: Lambda of the endosymbiont world. *Trends Microbiol.* **2010**, *18*, 173–181. [CrossRef] [PubMed]

21. Bordenstein, S.R.; Bordenstein, S.R. Eukaryotic association module in phage WO genomes from *Wolbachia*. *Nat. Commun.* **2016**, *7*, 13155. [CrossRef] [PubMed]

22. Pan, X.; Lührmann, A.; Satoh, A.; Laskowski-Arce, M.A.; Roy, C.R. Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* **2008**, *320*, 1651–1654. [CrossRef] [PubMed]

23. Richards, A.M.; Von Dwingelo, J.E.; Price, C.T.; Abu Kwaik, Y. Cellular microbiology and molecular ecology of *Legionella*-amoeba interaction. *Virulence* **2013**, *4*, 307–314. [CrossRef] [PubMed]

24. Lührmann, A.; Nogueira, C.V.; Carey, K.L.; Roy, C.R. Inhibition of pathogen-induced apoptosis by a *Coxiella burnetii* type IV effector protein. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18997–19001. [CrossRef] [PubMed]

25. Rikihisa, Y.; Lin, M. Anaplasma phagocytophilum and *Ehrlichia chaffeensis* type IV secretion and ANK proteins. *Curr. Opin. Microbiol.* **2010**, *13*, 59–66. [CrossRef] [PubMed]

26. Papafotiou, G.; Oehler, S.; Savakis, C.; Bourtzis, K. Regulation of *Wolbachia* ankyrin domain encoding genes in *Drosophila* gonads. *Res. Microbiol.* **2011**, *162*, 764–772. [CrossRef] [PubMed]

27. Duron, O.; Boureux, A.; Echaubard, P.; Berthomieu, A.; Berticat, C.; Fort, P.; Weill, M. Variability and expression of ankyrin domain genes in *Wolbachia* variants infecting the mosquito *Culex pipiens*. *J. Bacteriol.* **2007**, *189*, 4442–4448. [CrossRef] [PubMed]

28. Sinkins, S.P.; Walker, T.; Lynd, A.R.; Steven, A.R.; Makepeace, B.L.; Godfray, H.C.J.; Parkhill, J. *Wolbachia* variability and host effects on crossing type in Culex mosquitoes. *Nature* **2005**, *436*, 257–260. [CrossRef] [PubMed]

29. Walker, T.; Klasson, L.; Sebaihia, M.; Sanders, M.J.; Thomson, N.R.; Parkhill, J.; Sinkins, S.P. Ankyrin repeat domain-encoding genes in the wPip strain of *Wolbachia* from the *Culex pipiens* group. *BMC Biol.* **2007**, *5*, 39. [CrossRef] [PubMed]

30. Pichon, S.; Bouchon, D.; Liu, C.; Chen, L.; Garrett, R.A.; Grève, P. The expression of one ankyrin PK2 allele of the WO prophage is correlated with the *Wolbachia* feminizing effect in isopods. *BMC Microbiol.* **2012**, *12*, 55. [CrossRef] [PubMed]

31. Pichon, S.; Bouchon, D.; Cordaux, R.; Chen, L.; Garrett, R.A.; Grève, P. Conservation of the Type IV secretion system throughout *Wolbachia* evolution. *Biochem. Biophys. Res. Commun.* **2009**, *385*, 557–562. [CrossRef] [PubMed]

32. Wu, M.; Sun, L.V.; Vamathevan, J.; Riegler, M.; Deboy, R.; Brownlie, J.C.; McGraw, E.A.; Martin, W.; Esser, C.; Ahmadinejad, N.; et al. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: A streamlined genome overrun by mobile genetic elements. *PLoS Biol.* **2004**, *2*, e69. [CrossRef] [PubMed]

33. Foster, J.; Ganatra, M.; Kamal, I.; Ware, J.; Makarova, K.; Ivanova, N.; Bhattacharyya, A.; Kapatral, V.; Kumar, S.; Posfai, J.; et al. The *Wolbachia* genome of *Brugia malayi*: Endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* **2005**, *3*. [CrossRef] [PubMed]

34. Félix, C.; Pichon, S.; Braquart-Varnier, C.; Braig, H.; Chen, L.; Garrett, R.A.; Martin, G.; Grève, P. Characterization and transcriptional analysis of two gene clusters for type IV secretion machinery in *Wolbachia* of *Armadillidium vulgare*. *Res. Microbiol.* **2008**, *159*, 481–485. [CrossRef] [PubMed]

35. Beckmann, J.F.; Ronau, J.A.; Hochstrasser, M. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* **2017**, *2*, 17007. [CrossRef] [PubMed]

36. LePage, D.P.; Metcalf, J.A.; Bordenstein, S.R.; On, J.; Perlmutter, J.I.; Shropshire, J.D.; Layton, E.M.; Funkhouser-Jones, L.J.; Beckmann, J.F.; Bordenstein, S.R. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* **2017**, *543*, 243–247. [CrossRef] [PubMed]

37. Bonneau, M.; Atyame, C.; Beji, M.; Justy, F.; Cohen-Gonsaud, M.; Sicard, M.; Weill, M. *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat. Commun.* **2018**, *9*, 319. [CrossRef] [PubMed]

38. Rice, D.W.; Sheehan, K.B.; Newton, I.L.G. Large-Scale Identification of *Wolbachia pipientis* Effectors. *Genome Biol. Evol.* **2017**, *9*, 1925–1937. [CrossRef] [PubMed]

39. Moret, Y.; Juchault, P.; Rigaud, T. *Wolbachia* endosymbiont responsible for cytoplasmic incompatibility in a terrestrial crustacean: Effects in natural and foreign hosts. *Heredity* **2001**, *86*, 325–332. [CrossRef] [PubMed]

40. Cordaux, R.; Pichon, S.; Hatira, H.B.A.; Doublet, V.; Grève, P.; Marcadé, I.; Braquart-Varnier, C.; Souty-Grosset, C.; Charfi-Cheikhrouha, F.; Bouchon, D. Widespread *Wolbachia* infection in terrestrial isopods and other crustaceans. *ZooKeys* **2012**, 123–131. [CrossRef] [PubMed]

41. Badawi, M.; Grève, P.; Cordaux, R. Feminization of the isopod *Cylisticus convexus* after transinfection of the wVulC *Wolbachia* Strain of *Armadillidium vulgare*. *PLoS ONE* **2015**, *10*, e0128660. [CrossRef] [PubMed]

42. Bouchon, D.; Rigaud, T.; Juchault, P. Evidence for widespread *Wolbachia* infection in isopod crustaceans: Molecular identification and host feminization. *Proc. Biol. Sci.* **1998**, *265*, 1081–1090. [CrossRef] [PubMed]

43. Le Clec'h, W.; Braquart-Varnier, C.; Raimond, M.; Ferdy, J.-B.; Bouchon, D.; Sicard, M. High virulence of *Wolbachia* after host switching: When autophagy hurts. *PLoS Pathog.* **2012**, *8*, e1002844. [CrossRef] [PubMed]

44. Chandler, C.H.; Badawi, M.; Moumen, B.; Grève, P.; Cordaux, R. Multiple Conserved Heteroplasmic Sites in tRNA genes in the mitochondrial genomes of terrestrial isopods (Oniscidea). *Genes Genomes Genet.* **2015**, *5*, 1317–1322. [CrossRef] [PubMed]

45. Darling, A.E.; Mau, B.; Perna, N.T. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **2010**, *5*, e11147. [CrossRef] [PubMed]

46. Husemann, P.; Stoye, J. r2cat: Synteny plots and comparative assembly. *Bioinformatics* **2010**, *26*, 570–571. [CrossRef] [PubMed]

47. Kent, W.J. BLAT—The BLAST-like alignment tool. *Genome Res.* **2002**, *12*, 656–664. [CrossRef] [PubMed]

48. Seemann, T. PROKKA: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]

49. Li, L.; Stoeckert, C.J.; Roos, D.S. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **2003**, *13*, 2178–2189. [CrossRef] [PubMed]

50. Suzuki, S. Androgenic Gland Hormone Is a Sex-Reversing Factor but Cannot Be a Sex-Determining Factor in the Female Crustacean Isopods *Armadillidium vulgare*. *Gen. Comp. Endocrinol.* **1999**, *115*, 370–378. [CrossRef] [PubMed]

51. Suzuki, S.; Yamasaki, K. Sexual bipotentiality of developing ovaries in the terrestrial isopod *Armadillidium vulgare* (Malacostraca, Crustacea). *Gen. Comp. Endocrinol.* **1997**, *107*, 136–146. [CrossRef] [PubMed]

52. Suzuki, S.; Yamasaki, K. Sex reversal by implantations of ethanol-treated androgenic glands of female isopods, *Armadillidium vulgare* (Malacostraca, Crustacea). *Gen. Comp. Endocrinol.* **1998**, *111*, 367–375. [CrossRef] [PubMed]

53. Pfaffl, M.W.; Tichopad, A.; Prgomet, C.; Neuvians, T.P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **2004**, *26*, 509–515. [CrossRef] [PubMed]

54. Zhao, S.; Fernald, R.D. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J. Comput. Biol.* **2005**, *12*, 1047–1064. [CrossRef] [PubMed]

55. R Development CORE TEAM. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008.

56. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [CrossRef] [PubMed]

57. Letunic, I.; Doerks, T.; Bork, P. SMART 7: Recent updates to the protein domain annotation resource. *Nucleic Acids Res.* **2012**, *40*, D302–D305. [CrossRef] [PubMed]

58. Marchler-Bauer, A.; Zheng, C.; Chitsaz, F.; Derbyshire, M.K.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Lanczycki, C.J.; et al. CDD: Conserved domains and protein three-dimensional structure. *Nucleic Acids Res.* **2013**, *41*, D348–D352. [CrossRef] [PubMed]

59. Tatusov, R.L.; Galperin, M.Y.; Natale, D.A.; Koonin, E.V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [CrossRef] [PubMed]

60. Tatusov, R.L.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Kiryutin, B.; Koonin, E.V.; Krylov, D.M.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; et al. The COG database: An updated version includes eukaryotes. *BMC Bioinform.* **2003**, *4*, 41. [CrossRef] [PubMed]

61. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8*, 785–786. [CrossRef] [PubMed]

62. Rubtsov, A.M.; Lopina, O.D. Ankyrins. *FEBS Lett.* **2000**, *482*, 1–5. [CrossRef]

63. Mosavi, L.K.; Cammett, T.J.; Desrosiers, D.C.; Peng, Z.-Y. The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci.* **2004**, *13*, 1435–1448. [CrossRef] [PubMed]

64. Kajava, A.V. Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.* **1998**, *277*, 519–527. [CrossRef] [PubMed]

65. Kobe, B.; Kajava, A.V. The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* **2001**, *11*, 725–732. [CrossRef]

66. Blatch, G.L.; Lässle, M. The tetratricopeptide repeat: A structural motif mediating protein-protein interactions. *BioEssays* **1999**, *21*, 932–939. [CrossRef]

67. Walenta, J.H.; Didier, A.J.; Liu, X.; Krämer, H. The Golgi-Associated hook3 protein is a member of a novel family of microtubule-binding proteins. *J. Cell Biol.* **2001**, *152*, 923–934. [CrossRef] [PubMed]

68. Kerres, A.; Vietmeier-Decker, C.; Ortiz, J.; Karig, I.; Beuter, C.; Hegemann, J.; Lechner, J.; Fleig, U. The fission yeast kinetochore component Spc7 associates with the EB1 family member Mal3 and is required for kinetochore-spindle association. *Mol. Biol. Cell* **2004**, *15*, 5255–5267. [CrossRef] [PubMed]

69. Hirano, M.; Hirano, T. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. *EMBO J.* **2002**, *21*, 5733–5744. [CrossRef] [PubMed]

70. Jensen, R.B.; Shapiro, L. Cell-cycle-regulated expression and subcellular localization of the *Caulobacter crescentus* SMC Chromosome Structural Protein. *J. Bacteriol.* **2003**, *185*, 3068–3075. [CrossRef] [PubMed]

71. Yu, R.C.; Hanson, P.I.; Jahn, R.; Brünger, A.T. Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat. Struct. Biol.* **1998**, *5*, 803–811. [CrossRef] [PubMed]

72. Gorina, S.; Pavletich, N.P. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science* **1996**, *274*, 1001–1005. [CrossRef] [PubMed]

73. Anthonsen, H.W.; Baptista, A.; Drabløs, F.; Martel, P.; Petersen, S.B.; Sebastião, M.; Vaz, L. Lipases and esterases: A review of their sequences, structure and evolution. *Biotechnol. Annu. Rev.* **1995**, *1*, 315–371. [PubMed]

74. Wong, H.; Schotz, M.C. The lipase gene family. *J. Lipid Res.* **2002**, *43*, 993–999. [CrossRef] [PubMed]

75. Nocek, B.P.; Gillner, D.M.; Fan, Y.; Holz, R.C.; Joachimiak, A. Structural basis for catalysis by the mono and dimetalated forms of the dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase. *J. Mol. Biol.* **2010**, *397*, 617–626. [CrossRef] [PubMed]

76. Rowsell, S.; Pauptit, R.A.; Tucker, A.D.; Melton, R.G.; Blow, D.M.; Brick, P. Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure* **1997**, *5*, 337–347. [CrossRef]

77. John, G.B.; Shang, Y.; Li, L.; Renken, C.; Mannella, C.A.; Selker, J.M.L.; Rangell, L.; Bennett, M.J.; Zha, J. The mitochondrial inner membrane protein mitofilin controls cristae morphology. *Mol. Biol. Cell* **2005**, *16*, 1543–1554. [CrossRef] [PubMed]

78. Odgren, P.R.; Toukatly, G.; Bangs, P.L.; Gilmore, R.; Fey, E.G. Molecular characterization of mitofilin (HMP), a mitochondria-associated protein with predicted coiled coil and intermembrane space targeting domains. *J. Cell Sci.* **1996**, *109 Pt 9*, 2253–2264. [PubMed]

79. Fields, K.A.; Fischer, E.; Hackstadt, T. Inhibition of fusion of *Chlamydia trachomatis* inclusions at 32 degrees C correlates with restricted export of IncA. *Infect. Immun.* **2002**, *70*, 3816–3823. [CrossRef] [PubMed]

80. Pautsch, A.; Schulz, G.E. Structure of the outer membrane protein a transmembrane domain. *Nat. Struct. Biol.* **1998**, *5*, 1013–1017. [CrossRef] [PubMed]

81. Islam, M.S. (Ed.) *Transient Receptor Potential Channels*; Springer: Dordrecht, The Netherlands, 2011; ISBN 978-94-007-0264-6.

82. Gutzwiller, F.; Carmo, C.R.; Miller, D.E.; Rice, D.W.; Newton, I.L.G.; Hawley, R.S.; Teixeira, L.; Bergman, C.M. Dynamics of *Wolbachia pipientis* gene expression across the *Drosophila melanogaster* life cycle. *Genes Genomes Genet.* **2015**, *5*, 2843–2856. [CrossRef] [PubMed]

83. Dossi, F.C.A.; da Silva, E.P.; Cônsoli, F.L. Population dynamics and growth rates of endosymbionts during *Diaphorina citri* (Hemiptera, Liviidae) ontogeny. *Microb. Ecol.* **2014**, *68*, 881–889. [CrossRef] [PubMed]

84. Fast, E.M.; Toomey, M.E.; Panaram, K.; Desjardins, D.; Kolaczyk, E.D.; Frydman, H.M. *Wolbachia* enhance *Drosophila* stem cell proliferation and target the germline stem cell niche. *Science* **2011**, *334*, 990–992. [CrossRef] [PubMed]

85. Fenn, K.; Blaxter, M. Quantification of *Wolbachia* bacteria in *Brugia malayi* through the nematode lifecycle. *Mol. Biochem. Parasitol.* **2004**, *137*, 361–364. [CrossRef] [PubMed]

86. Landmann, F.; Foster, J.M.; Slatko, B.; Sullivan, W. Asymmetric *Wolbachia* segregation during early *Brugia malayi* embryogenesis determines its distribution in adult host tissues. *PLoS Negl. Trop. Dis.* **2010**, *4*, e758. [CrossRef] [PubMed]

87. Landmann, F.; Bain, O.; Martin, C.; Uni, S.; Taylor, M.J.; Sullivan, W. Both asymmetric mitotic segregation and cell-to-cell invasion are required for stable germline transmission of *Wolbachia* in filarial nematodes. *Biol. Open* **2012**, *1*, 536–547. [CrossRef] [PubMed]

88. Landmann, F.; Foster, J.M.; Michalski, M.L.; Slatko, B.E.; Sullivan, W. Co-evolution between an endosymbiont and its nematode host: *Wolbachia* asymmetric posterior localization and AP polarity establishment. *PLoS Negl. Trop. Dis.* **2014**, *8*, e3096. [CrossRef] [PubMed]

89.   Creasey, E.A.; Isberg, R.R. Maintenance of vacuole integrity by bacterial pathogens. *Curr. Opin. Microbiol.* **2014**, *17*, 46–52. [CrossRef] [PubMed]

90.   Kumar, Y.; Valdivia, R.H. Leading a sheltered life: Intracellular pathogens and maintenance of vacuolar compartments. *Cell Host Microbe* **2009**, *5*, 593–601. [CrossRef] [PubMed]

91.   Juchault, P.; Legrand, J.-J. Modification de la sex ratio dans les croisements entre différentes populations du crustacé oniscoïde *Armadllidium vulgare* Latr. Notion de déterminisme polygénique et épigénétique du sexe. *Arch. Zool. Exp. Gen.* **1976**, *1117*, 81–93. (In French)

92.   Juchault, P.; Legrand, J.-J.; Martin, G. Action interspécifique du facteur épigénétique féminisant responsable de la thélygénie et de l'intersexualité du crustacé *Armadillididium vulgare* (Isopode Oniscoïde). *Ann. Embryol. Morphog.* **1974**, *7*, 265–276. (In French)

93.   Juchault, P.; Legrand, J.-J. Nature et action interspécifique du facteur épigénétique féminisant responsable d'une perturbation totale ou partielle de l'équilibre endocrinien contrôlant le phénotype sexuel du crustacé *Armadillidium vulgare* (Isopode Oniscoïde). *Ann. Endocrinol.* **1974**, *35*, 387–392. (In French)

94.   Becking, T.; Giraud, I.; Raimond, M.; Moumen, B.; Chandler, C.; Cordaux, R.; Gilbert, C. Diversity and evolution of sex determination systems in terrestrial isopods. *Sci. Rep.* **2017**, *7*, 1084. [CrossRef] [PubMed]