

© Springer-Verlag New York Inc. 2000

# **Distribution and Evolution of Bacteriophage WO in** *Wolbachia,* **the Endosymbiont Causing Sexual Alterations in Arthropods**

**Shinji Masui, Satoru Kamoda, Tetsuhiko Sasaki, Hajime Ishikawa**

Department of Biological Sciences, Graduate School of Science, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received: 28 January 2000 / Accepted: 3 August 2000

**Abstract.** *Wolbachia* are obligatory intracellular and maternally inherited bacteria, known to infect many species of arthropod. In this study, we discovered a bacteriophage-like genetic element in *Wolbachia,* which was tentatively named bacteriophage WO. The phylogenetic tree based on phage WO genes of several *Wolbachia* strains was not congruent with that based on chromosomal genes of the same strains, suggesting that phage WO was active and horizontally transmitted among various *Wolbachia* strains. All the strains of *Wolbachia* used in this study were infected with phage WO. Although the phage genome contained genes of diverse origins, the average G+C content and codon usage of these genes were quite similar to those of a chromosomal gene of *Wolbachia.* These results raised the possibility that phage WO has been associated with *Wolbachia* for a very long time, conferring some benefit to its hosts. The evolution and possible roles of phage WO in various reproductive alterations of insects caused by *Wolbachia* are discussed.

**Key words:** *Wolbachia* — *Rickettsiaceae* — Bacteriophage WO — Virulence — Ankyrin repeat — Phage evolution

# **Introduction**

*Wolbachia,* a member of the *Rickettsiaceae* family, are obligatory intracellular and maternally inherited bacteria known to infect a wide range of arthropods. A recent survey suggested that about 16% of all insect species are infected with *Wolbachia,* making it one of the most ubiquitous endosymbionts described to date (Werren et al. 1995). These bacteria cause a number of reproductive distortions in their hosts, including male-killing (Hurst et al. 1999), feminization of geneitc male (Rousset et al. 1992), parthenogenesis (Stouthamer et al. 1993), and, most commonly, cytoplasmic incompatibility (CI). CI expression usually results in embryonic death in crosses in which the male insect parent is infected with a *Wolbachia* strain and the female parent is either uninfected or infected with a different *Wolbachia* strain (Werren 1997). It appears likely that a *Wolbachia* strain is able to imprint the sperm of insects by an unknown mechanism and that this imprinting is undone only in eggs that are infected with the same *Wolbachia* strain. Molecular mechanisms and genes of *Wolbachia* that are responsible for these phenomena are still unknown.

There is sound evidence that virulence determinants of pathogenic bacteria are encoded in their bacteriophages. The examples include diphtheria toxin produced by *Corynebacterium diphtheriae* (Barksdale and Arden 1974; Laird and Groman 1976), staphylokinase and enterotoxin A of *Strapylococcus aureus* (Coleman et al. 1989), Shiga-like toxin and enterohemolysin (Ehly2) produced by *Escherichia coli* (Newland et al. 1985; Beutin et al. 1993), and the neurotoxin of *Clostridium botulinum* (Barksdale and Arden 1974). Neither plasmids nor

GenBank accession numbers: AB035513 (*wsp, w*Cep), AB036665 (phage WO, *w*Tai), AB036666 (phage WO, *w*Kue), AB036649– AB036664 (*orf7* genes)

*Correspondence to:* Hajime Ishikawa; e-mail: iskw@biol.s.utokyo.ac.jp

#### Table 1. Insect and bacterial species used<sup>a</sup>



<sup>a</sup> *Drosophila simulans* Coffs Harbour S-20 and three strains of *Orientia (Rickettsia) tsutsugamushi* were provided by S.L. O'Neill and H. Urakami, respectively. The moth, *Ostrinia scapulalis,* infected with *w*Sca, expresses feminization (Hoshizaki, personal communication). Transinfected strains of *E. kuehniella,* carrying *w*CauA, *w*CauB, and *w*Sca, were established by microinjecting the corresponding *Wolbachia* strains into the eggs of tetracycline-treated strains of *E. kuehniella* (Sasaki, unpublished). *T. taiwanemma* (tet) is an uninfected strain generated by tetracycline treatment (Masui et al. 1997)

bacteriophages had been detected in the *Rickettsiaceae* family to which *Wolbachia* belongs. In the present study, we identified a bacteriophage-like genetic element of *Wolbachia* for the first time, which was tentatively named bacteriophage WO. The phage was detected in all the *Wolbachia* strains tested, and its genome displayed a history of extensive horizontal gene transfer.

# **Materials and Methods**

### *Materials*

Insects and bacterial strains used in this study are listed in Table 1. DNA and RNA were extracted from these materials by a standard method (Sambrook et al. 1989).

# *Library Construction, Screening, and Sequencing*

Total DNA of *Ephestia kuehniella* infected with *Wolbachia w*Kue was partially digested with *Sau*3AI (TAKARA), and a DNA library was constructed using the Lambda DASH II *Bam*HI Vector kit (Stratagene) and *E. coli* XLI-Blue MRA (P2) as a host strain. The DNA library of *Teleogryllus taiwanemma,* containing clones of DNA from *Wolbachia* strain *w*Tai, was used to isolate a clone containing ISW*l,* the insertion sequence of *Wolbachia* (Masui et al. 1999). Insert DNAs of clones were PCR amplified and their nucleotide sequences were determined by a primer walking method using an ABI PRISM 310 Genetic Analyzer.

#### *Southern Blotting*

Total DNA was digested to completion with *Eco*RI, electrophoresed on 1% agarose gels, and transferred to a nylon membrane (Hybond-N+; Amersham Pharmacia Biotech). Prehybridization and hybridization were performed according to the manufacturer's recommendations. The final wash was done using  $2 \times$  SSC at 55 $\degree$ C.

#### *Probes*

For the ISW*l* probe, the insert DNA of plasmid pISW*l* (Masui et al. 1999) was employed. The region encompassing nucleotide positions 7353–7761 of phage WO in *w*Tai (corresponding to the positions of the sequence deposited in GenBank under accession number AB035515) were PCR amplified with the primers phgWOF (5'-CCCACATG-AGCCAATGACGTCTG-3') and phgWOR (5'-CGTTCGCTCTG-CAAGTAACTCCATTAAAAC-3'), and the PCR product was used as an *orf*7 probe. These probes were labeled with  $[\alpha^{-32}P]$ dCTP using the *Bca*BEST labeling kit (Takara) and put in a hybridization solution.

# *RT-PCR*

Total RNA from *T. taiwanemma* and *E. kuehniella,* which harbor *w*Tai and *w*Kue, respectively, were reverse transcribed by SuperScript II RT (Gibco BRL) using the primer phgWOorf7RT  $(5'$ -TATTGCACT-CAGTATCTCTG-3'), PCR was performed using the primers phgWOR and phgWOR.

#### *Phylogenetic Analysis*

The *orf7* and *wsp* genes were PCR amplified with sets of primers phgWOF–phgWOR and wsp81F–wsp691R (Zhou et al. 1998), respectively. The PCR conditions for the set of primers phgWOF–phgWOR were 35 cycles of 30 s at 94°C, 30 s at 57°C and 1 min at 65°C. All the PCR products were cloned into pCRII-TOPO using the TOPOTA cloning kit (Invitrogen) and sequenced with T7 and SP6 primers. Clustal W was used to align sequences with manual refinement and to construct a neighbor-joining (NJ) tree. PAUP version 3.1.1 was used for maximum-parsimony (MP) analysis.

# *DNA Sequence Analysis*

DNA sequences were analyzed using DNASIS-MAC (Version 3.7; Hitachi). The predicted coding regions were defined as ORFs that consist of more than 100 codons and start with ATG, GTG, and TTG. Homology searches were performed using the BLAST programs at the National Center for Biotechnology Information (NCBI) server (Bethesda, MD).

# **Results**

# *Cloning and Identification of a Bacteriophage-Like Sequence in* Wolbachia

We previously reported the identification of the insertion sequence ISW*l* and the sequence analysis of its flanking regions in *Wolbachia* strain *w*Tai (Masui et al. 1999). In the course of the study, a clone that contained a gene



B

 $\overline{1}$ 



**Fig. 1. A** A bacteriophage-like element of *Wolbachia w*Tai. *Numbers in circles and parentheses* denote the percentage identity of predicted amino acid (aa) sequences between the indicated ORFs. Orf3 was similar to *E. coli* hypothetical protein O375, and a transposon Tn5-like sequence was observed in an opposite orientation between *orf4* and *orf6.* The partial sequence of Orf1 had 22% aa sequence identity to



mouse ankyrin, a membrane protein included in the cytoskelton. **B** Alignments of the predicted aa sequences and their  $\lambda$  homologues. GPA (the accession number of the Entrez protein search at the NCBI is AAA96534), GPB (AAA96536), and GPC (AAA96537) represent the products of genes A, B, and C, respectively. *Asterisks* indicate identical amino acids. *Numbers* show the amino acid positions of each proteins.

similar to the terminase large subunit gene of bacteriophage  $\lambda$  in the vicinity of ISW*l* was obtained from the DNA library of *T. taiwanemma* infected by *Wolbachia* strain *w*Tai. In addition, sequence analysis of this clone revealed that there were ORFs with a high sequence similarity to genes B and C of phage  $\lambda$  in the vicinity of the terminase-like gene (Figs. 1A and B). Although other ORFs in this region had no similarity to any phage genes, we assumed that all these ORFs constituted a prophagelike element, which we tentatively named phage WO.



**Fig. 2. A** Southern hybridization of *Wolbachia* and *Rickettsia* DNA with the *orf7* sequence from *Wolbachia w*Tai. tet, *T. taiwanemma* (tet); Rt, *O. tsutsugamushi* LF1. For other abbreviations, see Table 1. No signal was detected in *O. tsutsugamushi* LA1 or Gilliam (data not shown). **B** RT-PCR of *orf7.* RT + and − indicate the presence and absence of reverse transciptase in the reaction, respectively. P is the positive control for the PCR reaction using the total DNA of *T. taiwanemma* infected by *w*Tai as a template. N is the negative control using no template DNA.

# *Distribution and Expression of Phage WO Genes*

Southern blotting was performed using *orf7* of phage WO as a probe (Fig. 2A). As a result, this gene was observed in all eight *Wolbachia* strains tested, though not in *Orientia tsutsugamushi,* a close relative of *Wolbachia,* which is a presumed sex ratio distorter in mites (Takahashi et al. 1997). The hybridization signals detected were highly variable in size and strength, indicating that various phage WO types were present in a single strain of *Wolbachia.* The expression of *orf7* in *w*Tai and *w*Kue was examined by RT-PCR using the primers phgWOF and phgWOR, and positive signals were obtained in both strains (Fig. 2B).

# *Phylogenetic Analysis of Phage WO*

Phylogenetic analysis using *ftsZ* encoding a bacterial cell-cycle protein had demonstrated that *Wolbachia* strains in insects were clearly divided into two major groups, designated A and B (Werren et al. 1995). Recently, the outer membrane protein gene *wsp* was used for finer-scale phylogeny (Zhou et al. 1998). In an effort to determine whether phage WO was transmitted horizontally between the strains of *Wolbachia,* phylogenetic analysis of *orf7* was performed using *wsp* to construct a control phylogenetic tree to compare with the *orf7* tree. The *orf7* sequences in various *Wolbachia* strains were PCR amplified, and the products were cloned and sequenced. It turned out that there were several different *orf7* sequences in single *Wolbachia* strains. NJ analysis of *wsp* and *orf7* sequences showed that there was no concordance between the topologies of the two trees (Fig. 3). MP analysis also resulted in the same tree discordance (data not shown). The three clades of the *orf7* tree appeared to include both the A and the B groupderived sequences, as shown in the squares in Fig. 3. *Wolbachia* strains *w*CauA and *w*CauB are usually harbored by the same moth, *E. cautella* (Werren et al. 1995; Sasaki and Ishikawa 1999). The *orf7* sequences *w*CauA-1 and *w*CauB-1 were identical, and *w*CauA-2 was closely related to *w*CauB-3, indicating that *Wolbachia* strains harbored by the same species tend to share closely related phage WO types. These results suggest that phage WO was transmitted horizontally among different strains of *Wolbachia.*

#### *Sequence Analysis of Phage WO*

In an attempt to understand the structure of the phage WO genome, a DNA library was constructed using DNA from *E. kuehniella* harboring the *w*Kue strain and probed by the *orf7* sequence of *w*Tai. As a result, four overlapping clones that encompassed 33 kb were obtained, and the complete nucleotide sequence of this region was determined. ORF search of this region detected 33 ORFs (Fig. 4), which were designated *gp1–gp33,* encoding predicted proteins Gp1–Gp33. The results of an homology search of these predicted proteins are shown in Table 2. It was suggested that these genes had diverse origins. While some capsid proteins were similar to those of phage  $\lambda$ , an integrase found in Gp31 belonged to the resolvase/invertase family, whose members were different from that integrase of phage  $\lambda$  with respect to amino acid (aa) sequence and mechanisms of catalysis. Base plate assemble proteins of phage WO were similar to those of phage P2, whereas Gp12 was similar to adeninespecific methyltransferase of phage PhiCh1. Gp5 was similar to RepA, the protein involved in replication regulation of plasmid RSF1010. Several proteins containing ankyrin-like repeats were observed, which had about 40% aa sequence identity to each other. While ankyrinlike repeats are widely distributed, from bacteria to mammals, with a variety of functions, the ankyrin-like repeats



**Fig. 3.** Phylogenetic trees based on genes of two origins. These NJ trees were calculated using Clustal W, excluding positions of insertions–deletions and with correction for multiple substitutions. *Numbers* next to nodes indicate bootstrapping probabilities of 1000. *Broken lines* show linkages between strains of *Wolbachia* and phage WO strains conveyed by them. *Outlined type* represents A-group strains of *Wolbachia.* In the *orf7* tree, different *orf7* sequences detected in a single

*Wolbachia* strain are numbered arbitrarily. *Squares* indicate the pair of phage WO types from both A- and B-group strains of *Wolbachia.* Accession numbers of *wsp* genes in GenBank: *w*Ri, AF020070; *w*Cof, AF020067; *w*CauA, AB024571; *w*CauB, AB024572; *w*Tai, AB035514; *w*Kue, AB024569. The sequence lengths of *orf7* and *wsp* used for this analysis were about 350 and 440 bp, respectively.



**Fig. 4.** The ORF map of the region containing phage WO genes of *w*Kue and *w*Tai. *Arrows* indicate ORFs and their orientations, with associated ORF numbers corresponding to those in Table 2. Patterns in arrows indicate predicted origins of those genes: *horizontal stripes,* a plasmid; *vertical stripes,* bacteriophages; *black,* eukaryotes. E, *Eco*RI sites.

found in the 33-kb region were more similar to those of mammals and plants than those of bacteria (Table 2). Although diverse origins were suggested for these genes, their G+C contents were in a narrow range (32.3– 38.8%), which is much the same as that of the *wsp* gene (37.0%). There was no significant difference in codon usage between these genes, i.e., they all had strong A/Tbiased codon usage (data not shown).

# **Discussion**

Although the presence of virus-like particles in *Wolbachia* had been predicted by electron microscopy (Wright et al. 1978), no phages were detected in the bacteria. In

the present study, we identified the phage-related genes in *Wolbachia* for the first time. This also represents the first report on such genes in the *Rickettsiaceae* family, whose members are all obligate intracellular parasites or symbionts. The phage gene was expressed (Fig. 2B), and phylogenetic analysis revealed frequent transmission of the gene (Fig. 3), suggesting that phage WO is active and transferred horizontally between different strains of *Wolbachia* that are harbored in the cytoplasm of different insect hosts. Although the boundary of the phage WO genome on the *Wolbachia* chromosome was not determined in this study, it appears likely that the region from *gp12* to *gp31* is, at least, a part of the phage WO genome, since many phage-related genes reside in this region (Fig. 4 and Table 2). Southern blotting using the phage





<sup>a</sup> Similar proteins were shown for the best matches by the BLAST2 search at the NCBI

gene gave numerous bands (Fig. 2), due to both the high variability of the phage WO genes and the presence of several types of phage WO in a single *Wolbachia* strain. Numerous bands were also observed only when Southern blotting was performed using a DNA fragment from the region between *gp5* and *gp31,* though Southern blotting using *wsp* gene or other chromosomal genes of *Wolbachia* usually gave single bands (data not shown). Taken together, it is conceivable that the phage WO genome corresponds to the 25 kb region extending from *gp5* to *gp31.*

Comparison among the complete genomic sequences of several phages has revealed that although they have similar genomic organization, the genomes are mosaic in nature, with regions of obvious sequence similarity interspersed with segments that are apparently unrelated (Campbell et al. 1994; Hendrix et al. 1999). This suggests that all dsDNA phage genomes are mosaics that exchange genes horizontally with a large common genetic pool (Hendrix et al. 1999). Phage WO genome also had mosaic origins in capsid, baseplate assembly proteins and integrase genes. Moreover, the phage WO genome contained a gene derived from a plasmid and genes encoding several ankyrin-like proteins similar to those of mammals and plants (Fig. 4). This raises the possibility that a number of genes of considerably diverse origins are included in the common genetic pool of dsDNA phages, out of which these genes are distributed to many bacteria species.

Bacteriophages code for a variety of toxin genes expressed by pathogenic bacteria. The persistence of these genes in the phage genomes may indicate that there is an evolutionary advantage for phages to carry such genes, perhaps due to enhanced replication of bacteria carrying these virulence determinants (Cheetham and Katz 1995). *Wolbachia* are detected all over the world and spreading rapidly, by causing various reproductive alterations to their hosts, through which they can efficiently spread in the host populations (Werren 1997). Our present study implies that most of *Wolbachia* carry phage WO (Fig. 2). In addition, the G+C content and codon usage of the phage WO genes were quite similar to those of a chromosomal gene of *Wolbachia,* and there were extreme sequence diversities in the phage WO genes (Figs. 2 and 3). All these findings suggest the ancient origin of phage WO. It is probable that phage WO has associated with *Wolbachia* for a very long time, enabling *Wolbachia* to remain in insect hosts by producing some virulence factors. Since ankyrin-like repeats are considered to be a motif for protein–protein interaction (Sedgwick and Smerdon 1999), it is possible that the protein containing eukaryotic ankyrin-like repeats found in the phage WO genome play a role in the reproductive alteration events of insect hosts through their ability to interact with other proteins.

It has been suggested that the ability of *Wolbachia* to spread in insect populations effectively could be used as a vehicle to drive desirable genotypes into wild insect populations, e.g., genes that prevent insect vectors from transmitting pathogens to humans, livestock, and plants (Sinkins et al. 1997). However, due to the lack of a genetic system to manipulate this bacterium, very little is known about molecular mechanisms that underlie the interaction of this agent with its host. Phage WO can hopefully be used as a vector to transform these fastidious bacteria without any culture step.

*Acknowledgments.* We thank Dr. S.L. O'Neill (Yale University) and Dr. H. Urakami (Niigata College of Pharmacy) for providing materials. We also thank Dr. S. Hoshizaki (University of Tokyo) for providing materials and information about *Wolbachia* infecting *O. scapulalis.* This study was supported in part by Research Fellowships from the Japan Society for the Promotion of Science for Young Scientists.

#### **References**

- Barksdale L, Arden SB (1974) Persisting bacteriophage infections, lysogeny, and phage conversions. Annu Rev Microbiol 28:265–299
- Beutin L, Stroeher UH, Manning PA (1993) Isolation of enterohaemolysin (Ehly2)-associated sequences encoded on temperate phages of *Escherichia coli.* Gene 132:95–99
- Campbell A (1994) Comparative molecular biology of lambdoid phages. Annu Rev Microbiol 48:193–222
- Cheetham BF, Katz ME (1995) A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. Mol Microbiol 18:201–208
- Coleman DC, Sullivan DJ, Russell RJ, Arbuthnott JP, Carey BF, Pomeroy HM (1989) *Straphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: Molecular mechanism of triple conversion. J Gen Microbiol 135:1679–1697
- Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF (1999) Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. Proc Natl Acad Sci USA 96: 2192–2197
- Hurst GD, Graf von der Schulenburg JH, Majerus TM, Bertrand D, Zakharov IA, Baungaard J, Volkl W, Stouthamer R, Majerus ME (1999) Invasion of one insect species, *Adalia bipunctata,* by two different male-killing bacteria. Insect Mol Biol 8:133–139
- Laird W, Groman N (1976) Orientation of the tox gene in the prophage of corynebacteriophage beta. J Virol 19:228–231
- Masui S, Sasaki T, Ishikawa H (1997) *groE*-homologous operon of *Wolbachia,* an intracellular symbiont of arthropods: A new approach for their phylogeny. Zool Sci 14:701–706
- Masui S, Kamoda S, Sasaki T, Ishikawa H (1999) The first detection of the insertion seqeunce ISW*l* in the intracellular reproductive parasite *Wolbachia.* Plasmid 42:13–19
- Newland JW, Strockbine NA, Miller SF, O'Brien AD, Holmes RK (1985) Cloning of Shiga-like toxin structural genes from a toxin converting phage of *Escherichia coli.* Science 230:179–181
- Rousset F, Bouchon D, Pintureau B, Juchault P, Solignac M (1992) *Wolbachia* endosymbionts responsible for various alterations of sexuality in arthropods. Proc R Soc Lond B 250:91–98
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning—A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Sasaki T, Ishikawa H (1999) *Wolbachia* infections and cytoplasmic incompatibility in the almond moth and the Mediterranean flour moth. Zool Sci 16:739–744
- Sedgwick SG, Smerdon SJ (1999) The ankyrin repeat: A diversity of interactions on a common structural framework. Trends Biochem Sci 24:311–316
- Sinkins SP, Curtis CF, O'Neill SL (1997) The potential application of inherited symbiont systems to pest control. In: O'Neill SL, Hoffmann AA, Werren JH (eds) Influential passengers: Inherited microorganisms and arthropod reproduction. Oxford University Press, Oxford, pp 155–175
- Stouthamer R, Breeuwert JA, Luck RF, Werren JH (1993) Molecular identification of microorganisms associated with parthenogenesis. Nature 361:66–68
- Takahashi M, Urakami H, Yoshida Y, Furuya Y, Misumi H, Hori E, Kawamura A Jr, Tanaka H (1997) Occurrence of high ratio of males after introduction of minocycline in a colony of *Leptotrombidium fletcheri* infected with *Orientia tsutsugamushi.* Eur J Epidemiol 13:79–86
- Werren JH (1997) Biology of *Wolbachia.* Annu Rev Entomol 42:587– 609
- Werren JH, Windsor D, Guo LR (1995) Distribution of *Wolbachia* among neotropical anthropods. Proc R Soc Lond B 262:197–204
- Wright JD, Sjostrand FS, Portaro JK, Barr AR (1978) The ultrastructure of the rickettsia-like microorganism *Wolbachia pipientis* and associated virus-like bodies in the mosquito *Culex pipiens.* J Ultrastruct Res 63:79–85
- Zhou W, Rousset F, O'Neill SL (1998) Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc R Soc Lond B Biol Sci 265:509–515