Letter to the Editor

How Many *Wolbachia* Supergroups Exist?

Nathan Lo,†* *Maurizio Casiraghi,** *Emanuela Salati,** *Chiara Bazzocchi,** *and Claudio Bandi**

*Dipartimento di Patologia Animale, Igiene e Sanità Pubblica Veterinaria, Sezione di Patologia Generale e Parassitologia, Università di Milano, Italy; †National Institute of Agrobiological Sciences, Tsukuba, Japan

Obligate intracellular bacteria of the genus *Wolbachia* (Class Alphaproteobacteria, Order Rickettsiales) are currently divided into four taxonomic supergroups on the basis of clustering patterns in *ftsZ*-based phylogenetic trees (Werren, Zhang, and Guo 1995; Bandi et al. 1998). Supergroups A and B are found only in arthropods, whereas C and D are found only in filarial nematodes. The term supergroup has recently been employed to avoid confusion with designation of more closely related groups based on *wsp* sequences (Zhou, Rousset, and O'Neill 1998). *Wolbachia* have generated substantial interest in recent years (Zimmer 2001), primarily because of the effects they have on their arthropod hosts, which include induction of cytoplasmic incompatibility (CI), parthenogenesis, feminization, and male-killing (reviewed in Stouthammer, Breeuwer, and Hurst 1999). Estimation of the phylogenetic relationships within each supergroup has provided useful information about the evolution and biology of these bacteria. The phylogenies of both A and B members have been found to be incongruent with that of their hosts, strongly suggesting horizontal transfer (Werren, Zhang, and Guo 1995). Recently, direct evidence for this phenomenon was found for parasitic wasps sharing a common food source (Huigens et al. 2000). Unlike the case of arthropods, the phylogeny of each nematode *Wolbachia* supergroup (C and D) appears to match that of their hosts (Casiraghi et al. 2001*a*), although further gene sequencing studies are required to confirm this. Such phylogenetic congruence would suggest a strictly dependent relationship, and this idea is supported by evidence that removal of *Wolbachia* using antibiotics has negative effects on the filariae they reside in (Bandi et al. 1999; Langworthy et al. 2000).

Progress toward answering several remaining questions about *Wolbachia* evolution—such as which of their host effects are primitive and which are derived, the type of animals they first invaded, and how they were transferred between arthropods and nematodes—is currently hindered by a poor understanding of the relationships between the supergroups. An improved estimate of *Wolbachia* phylogeny at this level will require: (1) the inclusion of sequence information from diverse, possibly as-yet-unknown taxa, (2) an appropriate choice of genes and outgroups, and (3) the use of sound data analysis techniques which enable statistical assessment of results. In this report, we demonstrate the presence of two divergent *Wolbachia* lineages, which increases the number of recognized supergroups from four to six. Phylogenetic analyses have been performed using likelihood-based techniques, which are becoming increasingly accepted as the most reliable and efficient for phylogenetic inference (Whelan, Liò, and Goldman 2001). Whereas some recent *Wolbachia* studies have explicitly used likelihood methods to examine within-supergroup relationships (Von der Schulenberg et al. 2000; Jiggins et al. 2001), to our knowledge this study represents the first likelihood-based analysis of relationships between supergroups.

A report by Vandekerchove et al. (1999) on endosymbionts from the springtail *Folsomia candida* (Collembola) proposed that these bacteria represent a fifth divergent lineage of *Wolbachia* (supergroup E). However, rather than using *ftsZ*—the gene originally used to describe supergroups A-D—the authors based their proposal on the relatively slow-evolving 16S rDNA. The evidence supporting the establishment of this new *Wolbachia* supergroup could be considered inconclusive for the following reasons: (1) the bootstrap value for the separation of the clade containing A + B members from the branch representing the *Wolbachia* from *F. candida* was marginal (55%), and (2) the divergence between the 16S rDNAs of *Wolbachia* from *F. candida* and those of A members was very low (~1%) (Vandekerchove et al. 1999).

In order to examine whether *Wolbachia* from *F. candida* do actually represent a divergent supergroup, we amplified and sequenced *ftsZ* from representatives of this strain. Additionally, *ftsZ* sequences from *Wolbachia* strains present in the two termite (Isoptera) species *Kalotermes flavicollis* and *Microcerotermes* sp. (Bandi et al. 1997) were obtained. *Wolbachia* from *K. flavicollis* had previously been shown to be divergent from other *Wolbachia* based on 16S rDNA sequences (Bandi et al. 1999). DNA was extracted from five pooled *F. candida* individuals and abdomen tissues from individual termites (Bandi et al. 1994). Initial PCR attempts with the commonly used primers ftsZf1 and ftsZr1 (Werren, Zhang, and Guo 1995) were met with failure, and thus two additional primers (ftsZunif 5′–3′ GG(CT)AA(AG) GGTGC(AG)GCAGAAGA; ftsZunir 5′–3′ ATC(AG) AT(AG)CCAGTTGCAAG) were designed based on an alignment of all arthropod and filarial *Wolbachia ftsZ* sequences. These primers enabled amplification of a 775-bp fragment (737 bp excluding primers) from *F. candida* and each termite specimen. Both strands of each fragment were sequenced directly using ABI technology.

Phylogenetic analyses were performed using Bayesian inference, maximum likelihood (ML), and maximum parsimony (MP) methods. The *ftsZ* sequences

Key words: *Wolbachia*, termite, *Folsomia*, likelihood, *ftsZ*, 16S rDNA.

Address for correspondence and reprints: Dipartimento di Patologia Animale, Ingiene e Sanità Pubblica Veterinaria, Sezione di Patologia Generale e Parassitologia Università di Milano, Via Celoria 10, Milano 20133, Italy. E-mail: nathanlo@affrc.go.jp.

Mol. Biol. Evol. 19(3):341–346. 2002

obtained in this study were aligned manually with those from 20 other *Wolbachia* from supergroups A-D. Of the 737 characters, 237 were variable and 178 parsimony informative. An appropriate model of sequence evolution (GTR + I + G) was selected via likelihood ratio tests (5% significance level) using the program Modeltest 3.06 (Posada and Crandall 1998). From calculations in TreePuzzle 5.0 (Strimmer and Von Haeseler 1996), no significant base composition heterogeneity was found among the 23 sequences. For Bayesian inference of phylogeny, the program MrBayes 2.01 (Huelsenbeck and Ronquist 2001) was used. Estimating parameters from the data, a total of 4,000 trees were obtained; the first 2,000 of these were considered the burn in and discarded. A 50% majority-rule consensus tree of the remaining 2,000 trees was produced. ML branch lengths for this consensus tree were calculated in PAUP*4.0b8 (Swofford 2000), estimating parameters from the data. Topologies were also estimated under ML criteria using the successive approximations method of Swofford et al. (1996) and by MP bootstrap analysis (1,000 replicates with 10 random addition replicates per bootstrap replicate) in PAUP*4.0b8.

Figure 1 shows the unrooted *ftsZ*-based topology of *Wolbachia* from arthropods and filarial nematodes, estimated using MrBayes 2.01. Trees estimated from ML and MP bootstrap analyses were identical to that in figure 1, with the exception that the levels of resolution within supergroup B were slightly different among the three topologies. Each of the previously described supergroups A, B, C, and D were found as divergent and coherent clusters, supported by high probability values from Bayesian inference (97%–100%), as well as high bootstrap values from MP (99%–100%). On the basis of the branching pattern of figure 1 and the ML branch lengths separating the four recognized supergroups, we propose the establishment of an additional two supergroups: one containing *Wolbachia* from *F. candida* (supergroup E, as originally suggested by Vandekerchove et al. [1999] based on 16S rDNA) and another containing *Wolbachia* from termites (supergroup F). To our knowledge, supergroups A and B have not been reported from either springtails or termites.

The topology of figure 1 suggested that neither the nematodes nor the arthropods *Wolbachia* are monophyletic. To check whether this result was significant, we calculated the ML score of the topology where supergroups C + D (and thus supergroups A + B + E + F) were constrained to be monophyletic and compared it to the ML score of figure 1 using the SH test (Shimodaira and Hasegawa 1999). The former ($\ln L = -3330.46$) was found to be only slightly lower than the latter ($\ln L = -3329.22$), and the two were not found to be significantly different ($P = 0.24$). The relatively short branches separating supergroups C, D, E, and F (indicated by the labels root 2 and root 5 in fig. 1) are suggestive of a rapid evolutionary radiation in the ancestors of these clades. An alternative explanation is that multiple hits in this gene have accumulated to the extent that the synapomorphies between certain supergroups have been erased.

In an attempt to examine which of the supergroups A-F are basal and which are derived, we calculated the ML values of topologies when the *ftsZ* of the outgroup *Anaplasma marginale* was constrained on branches leading to each supergroup, as well as on those separating two or more supergroups (see fig. 1, roots 1–9). Using the SH test to compare the nine different ML scores (which ranged from $-\ln L = 3901.46$ to 3902.65), no outgroup position was found to give a significantly higher likelihood score than any other (5% significance level). Thus, despite *ftsZ* from *A. marginale* being the closest known putative homologue of *Wolbachia ftsZ*, it does not enable resolution of the root of the *Wolbachia* tree. This is probably partly because of the large nucleotide distance in this gene between these two genera (~0.3 uncorrected substitutions per site; ~1.7 substitutions per site inferred from ML).

To check whether the relationships inferred from *ftsZ* were consistent with those from another gene, an analysis of 16S rDNA was performed. This gene was chosen firstly to re-examine the results of Vandekerchove et al. (1999) using likelihood-based methods and secondly because of the availability of sequences from diverse taxa, including *Wolbachia* from the weevil *Rhinocyllus conicus* (Campbell, Bragg, and Turner 1992) and the filarial nematode *Mansonella ozzardi* (Casiraghi et al. 2001*b*). To facilitate this analysis, near-full length sequences (1,312 bp) from *Wolbachia* present in the termites *K. flavicollis* and *Microcerotermes* sp. were amplified and sequenced as described in Bandi et al. (1994). Taking into account the secondary structure (Neefs et al. 1993), these were aligned with 20 previously obtained sequences. Of the 1,321 aligned characters, 281 were variable and 101 parsimony informative. The most appropriate model of sequence evolution for *Wolbachia* 16S rDNA chosen by Modeltest 3.06 was HKY+G. No evidence for significant base composition heterogeneity was found among the sequences. Figure 2 shows the topology for these sequences estimated using Bayesian inference. The topology from MP bootstrap analysis was identical to that in figure 2, and the ML topology differed only slightly with regard to the resolution within supergroups A and B. Overall, relationships were similar to those based on *ftsZ* (fig. 1), although branch lengths were considerably shorter, and supergroup A did not form a monophyletic group. This result shows that whereas 16S rDNA is useful for preliminary information about *Wolbachia* relationships, its rate of evolution appears unsuitable for providing clear resolution of the supergroups. Interestingly, the symbiont of *R. conicus* formed a clade with *Wolbachia* from termites. Although it appears to be closely related to supergroup F, *ftsZ* sequence information will be required to confirm this. The *Wolbachia* from *M. ozzardi* was found as a relatively long branch in the same clade as the strains from termites and *R. conicus*. This result is intriguing in the light of the question of whether filarial nematodes *Wolbachia* are a monophyletic assemblage. *ftsZ* sequence would obviously be useful to determine if the *Wolbachia* of *M. ozzardi* is a member of supergroup F

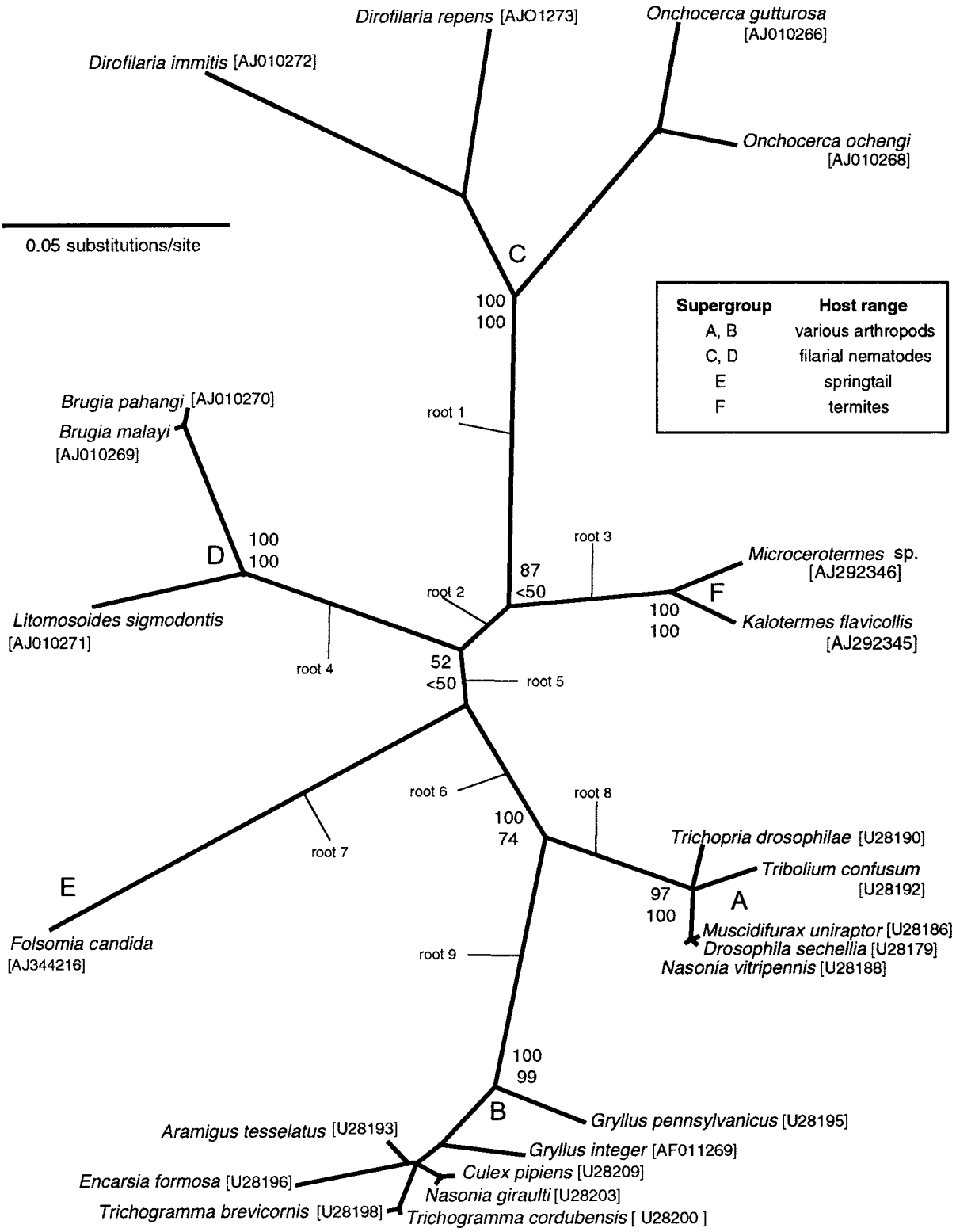

FIG. 1.—Unrooted phylogenetic tree of *Wolbachia* endosymbionts of arthropods and filarial nematodes based on *ftsZ*, estimated using Bayesian inference of phylogeny. Posterior probabilities supporting nodes of interest are shown above bootstrap values from a MP analysis. Names represent host species. Roots 1–9 indicate positions where the *ftsZ* gene of the outgroup *Anaplasma marginale* was constrained during likelihood estimations to examine the most appropriate root placement. Accession numbers are shown adjacent to each taxon. Each supergroup is labeled with one of the letters A-F. The existence of E and F was confirmed during this study.

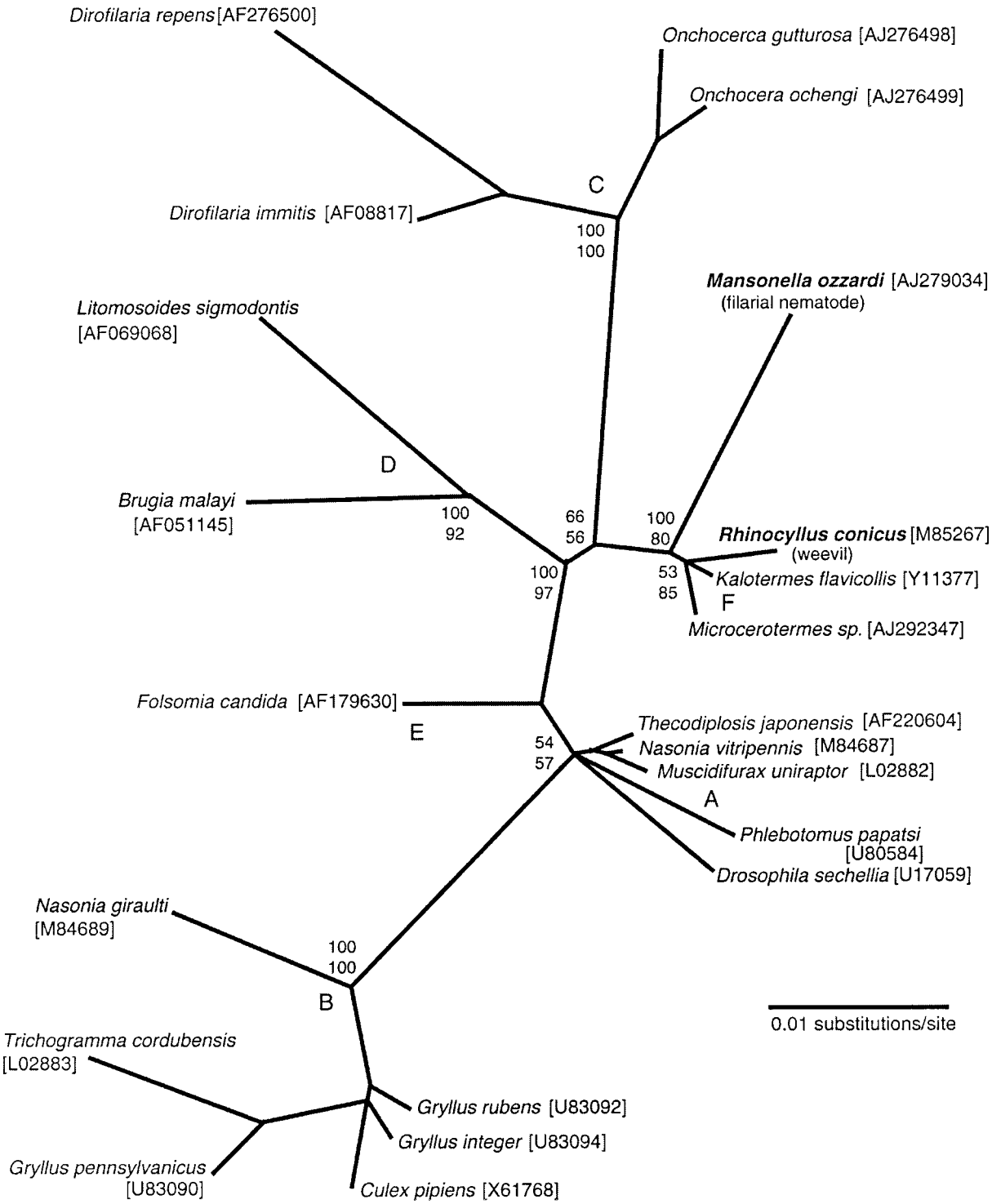

FIG. 2.—Unrooted phylogenetic tree of *Wolbachia* endosymbionts based on 16S rDNA sequences. Inference method, support values, and supergroup designations are as outlined in figure 1. Species names are those of the hosts. Names in bold indicate *Wolbachia* from taxa which have not yet been classified into supergroups on the basis of *ftsZ* sequences.

or whether it is a separate, divergent lineage. However, because of the difficulties in acquiring samples of this nematode, we have not yet been able to obtain this information. In a similar case to that of *ftsZ,* SH-tests using 16S rDNA from *A. marginale* and *Cowdria ruminatium* (another one of *Wolbachia*'s closest known relatives) showed that the position of the root in this tree is ambiguous when using this gene (data not shown).

In summary, we have confirmed the existence of two divergent lineages of *Wolbachia* by likelihood-based phylogenetic analysis of *ftsZ,* a gene that appears to be appropriate for the assessment of broader-level re-

lationships in this genus. Although less well resolved, the topology of a tree based on 16S rDNA from a similar group of taxa showed overall agreement with the tree inferred from *ftsZ*. Knowledge about the diversity of *Wolbachia* strains is desirable for a number of reasons. Firstly, future decisions on nomenclature within the genus will be aided by an understanding of the number of divergent strains and their differences at the DNA level. Secondly, detection of particular types of *Wolbachia* in arthropods (or other invertebrates) can be biased by the type of PCR primers used. Indeed, during this study, we were unable to obtain *ftsZ* amplification in *F. candida* and termites using the standard primers for arthropod A and B *Wolbachia* (ftsZf1 and ftsZr1; Werren, Zhang and Guo 1995). Incorporating sequence information from divergent *Wolbachia* during primer design might aid in the detection of *Wolbachia* in other arthropods as well as more diverse invertebrate hosts. Thirdly, an understanding of the phenotypic effects that the various divergent lineages have on their hosts, as well as their phylogenetic relationships, will provide clues on the nature of the ancestor to all *Wolbachia*, which is generally assumed to have been a CI-causing agent (Stouthammer, Breeuwer, and Hurst 1999, Pp. 75). The effects that E and F supergroup members have on their hosts is not yet clear, and further work on them should be encouraged. A fourth potential benefit of knowledge about *Wolbachia* diversity is related to the several genome sequencing projects that have been initiated in the genus (Bandi, Slatko, and O'Neill 1999). It has been revealed that the genome sizes of those from supergroups A and B (arthropods) are ~1.4–1.6 Mbp and are significantly larger than those of representatives from supergroups C and D (nematodes), which are 1.0–1.1 Mbp (Sun et al. 2001). Thus, determination of the genome size of *Wolbachia* from other lineages (such as those in termites and *F. candida*) followed by sequencing studies could provide answers to the question of whether there has been an overall reduction or increase in the genome size of *Wolbachia* since it first began invading invertebrate tissues.

Acknowledgments

Two anonymous reviewers provided useful suggestions. This work was supported by a MURST grant to Claudio Genchi. N.L. is supported by the Japanese Science and Technology Agency Fellowship Program.

LITERATURE CITED

BANDI, C., T. J. C. ANDERSON, C. GENCHI, and M. BLAXTER. 1998. Phylogeny of *Wolbachia* in filarial nematodes. Proc. R. Soc. Lond. B **265**:2407–2413.

BANDI, C., G. DAMIANI, L. MAGRASSI, A. GRIGOLO, R. FANI, and L. SACCHI. 1994. Flavobacteria as intracellular symbionts in cockroaches. Proc. R. Soc. Lond. B **257**:43–48.

BANDI, C., J. W. MCCALL, C. GENCHI, S. CORONA, L. VENCO, and L. SACCHI. 1999. Effects of tetracycline on the filarial worms *Brugia pahangi* and *Dirofilaria immitis* and their bacterial endosymbionts *Wolbachia*. Int. J. Parasitol. **29**: 357–364.

BANDI, C., M. SIRONI, C. A. NALEPA, S. CORONA, and L. SACCHI. 1997. Phylogenetically distant intracellular symbionts in termites. Parassitologia **39**:71–75.

BANDI, C., B. SLATKO, and S. L. O'NEILL. 1999. *Wolbachia* genomes and the many faces of symbiosis. Parasitol. Today **15**:428–429.

CAMPBELL, B. C., T. S. BRAGG, and C. E. TURNER. 1992. Phylogeny of symbiotic bacteria of four weevil species (Coleoptera: Curculionidae) based on analysis of 16S ribosomal DNA. Insect Biochem. Mol. Biol. **22**:415–421.

CASIRAGHI, M., T. J. C. ANDERSON, C. BANDI, C. BAZZOCCHI, and C. GENCHI. 2001*a*. A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of *Wolbachia* endosymbionts. Parasitology **122**:93–103.

CASIRAGHI, M., G. FAVIA, G. CANCRINI, A. BARTOLONI, and C. BANDI. 2001*b*. Molecular identification of *Wolbachia* from the filarial nematode *Mansonella ozzardi*. Parasitol. Res. **87**:417–420.

HUELSENBECK, J. P., and F. RONQUIST. 2001. MrBayes: bayesian inference of phylogeny. Bioinformatics **17**:754–755.

HUIGENS, M. E., R. F. LUCK, R. H. KLAASSEN, M. F. MAAS, M. J. TIMMERMANS, and R. STOUTHAMER. 2000. Infectious parthenogenesis. Nature **405**:178–179.

JIGGINS, F. M., J. H. VON DER SCHULENBERG, G. D. D. HURST, and M. E. N. MAJERUS. 2001. Recombination confounds interpretations of *Wolbachia* evolution. Proc. R. Soc. Lond. B **268**:1423–1427.

LANGWORTHY, N. G., A. RENZ, U. MACKENSTEDT, K. HENKLE-DUHRSEN, M. B. DE BRONSVOORT, V. N. TANYA, M. J. DONNELLY, and A. J. TREES. 2000. Macrofilaricidal activity of tetracycline against the filarial nematode *Onchocerca ochengi*: elimination of *Wolbachia* precedes worm death and suggests a dependent relationship. Proc. R. Soc. Lond. B **267**:1063–1069.

NEEFS, J., Y. VAN DE PEER, P. DE RIJK, S. CHAPELLE, and R. DE WACHTER. 1993. Compilation of small ribosomal subunit RNA structures. Nucleic Acids Res. **21**:3025–3049.

POSADA, D., and K. A. CRANDALL. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics **14**:817–818.

SHIMODAIRA, H., and M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. **16**:1114–1116.

STOUTHAMMER, R., J. A. J. BREEUWER, and G. D. D. HURST. 1999. *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. Ann. Rev. Microbiol. **53**:71–102.

STRIMMER, K., and A. VON HAESELER. 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13**:964–969.

SUN, L. V., J. FOSTER, G. TZERTZINIS, M. ONO, C. BANDI, B. E. SLATKO, and S. L. O'NEILL. 2001. Determination of *Wolbachia* genome size by pulse-field gel electrophoresis. J. Bacteriol. **183**:2219–2225.

SWOFFORD, D. L. 2000. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Mass.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. Hillis, C. Moritz, and B. K. Mable, eds. Molecular Systematics. 2nd edition. Sinauer Associates, Sunderland, Mass.

VANDEKERCHOVE, T. M. T., S. WATTEYNE, A. WILLEMS, J. G. SWINGS, J. MERTENS, and M. GILLIS. 1999. Phylogenetic analysis of the 16S rDNA of the cytoplasmic bacterium *Wolbachia* from the novel host *Folsomia candida* (Hexapoda, Collembola) and its implications for wolbachial taxonomy. FEMS Microbiol. Lett. **180**:279–286.

VON DER SCHULENBERG, J. H. G., G. D. D. HURST, T. M. E. HUIGENS, M. M. M. VAN MEER, F. M. JIGGINS, and M. E.

N. MAJERUS. 2000. Molecular evolution and phylogenetic utility of *Wolbachia ftsZ* and *wsp* gene sequences with special reference to the origin of male-killing. Mol. Biol. Evol. **17**:584–600.

WERREN, J. H., W. ZHANG, and L. R. GUO. 1995. Evolution and phylogeny of *Wolbachia*: reproductive parasites of arthropods. Proc. R. Soc. Lond. B **161**:55–63.

WHELAN, S., P. LIÒ, and N. GOLDMAN. 2001. Molecular phylogenetics: state-of-the-art methods for looking into the past. Trends Genet. **17**:262–272.

ZHOU, W., F. ROUSSET, and S. L. O'NEILL. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc. R. Soc. Lond. B. **265**:505–515.

ZIMMER, J. 2001. Meet the Herod bug. Nature **412**:12–14.

HOWARD OCHMAN, reviewing editor

Accepted October 8, 2001