Host-Symbiont Conflicts: Positive Selection on an Outer Membrane Protein of Parasitic but not Mutualistic Rickettsiaceae

Francis M. Jiggins, Gregory D. D. Hurst,† and Ziheng Yang†*

*Department of Genetics, University of Cambridge; and †Department of Biology, University College London

The Rickettsiaceae is a family of intracellular bacterial symbionts that includes both vertically transmitted parasites that spread by manipulating the reproduction of their host (*Wolbachia* in arthropods) and horizontally transmitted parasites (represented by *Cowdria ruminantium*), and mutualists (*Wolbachia pipientis* in nematode worms). We have investigated the nature of natural selection acting on an outer membrane protein, the *wsp* gene in *Wolbachia* and its homologue *map1* in *Cowdria*, thought likely to be involved in host-parasite interactions in these bacteria. The ratio of nonsynonymous to synonymous substitution rates (d_N/d_S) at individual amino acid sites or at lineages within the gene's phylogeny was estimated using maximum likelihood models of codon substitution. The first hypothesis we tested was that this protein is under positive selection in the parasitic but not in the mutualistic Rickettsiaceae. This hypothesis was supported as positive selection and was detected in *Cowdria* and arthropod *Wolbachia* sequence evolution but not in the evolution of *Wolbachia* sequences from nematodes. Furthermore, this selection was concentrated outside the transmembrane region of the protein and, therefore, in the regions of the protein that may interact with the host. The second hypothesis tested was that positive selection would be stronger in the strains of arthropod *Wolbachia* that distort the host sex ratio than in those that induce cytoplasmic incompatibility. However, we found no support for this hypothesis. In conclusion, our results are consistent with the hypothesis that antagonistic coevolution causes faster evolution of surface protein sequences in parasites than in mutualists. Confirmation of this conclusion awaits the replication of these results both in additional genes and across more bacterial taxa. The regions of the *wsp* and *map1* genes we identified as likely to be involved in host-parasite arms races should be examined in future studies of parasite virulence and host immune responses, and during the design of vaccines.

Introduction

Most animals and plants have intimate associations with a diverse range of microbial symbionts. These symbionts may be classified into mutualists, whose presence increases host fitness, and parasites, which are deleterious. When the association is parasitic, coevolution is expected to take the form of an arms race because host and parasite meet adaptation with counteradaptation (Ehrlich and Raven 1964). However, when the association is mutualistic, arms races driven by the host evolving to eliminate the pathogen are expected to be absent. This scenario generates the prediction that genes involved in host-symbiont interactions will evolve rapidly because of selection continually favoring novel traits, whereas those involved in mutualistic interactions will tend to be under stabilizing selection. The aim of this study was to test this hypothesis by directly estimating whether natural selection has caused the diversification of an outer membrane protein in parasites and mutualists.

The symbionts studied were intracellular bacterial symbionts of the family Rickettsiaceae. Most of the species in this family are arthropod-vectored parasites of vertebrates, which are represented in our study by the species *Cowdria ruminantium*, a tick-borne pathogen of ruminants. The second host-symbiont interaction seen in the Rickettsiaceae is reproductive parasitism, which is exhibited by *Wolbachia* bacteria infecting a range of arthropod hosts. These bacteria are primarily vertically transmitted from mother to offspring, and they spread by manipulating the reproduction of their hosts to enhance their own transmission. The main manipulations are distorting the host sex ratio toward females and inducing cytoplasmic incompatibility (CI) (note that conflict may be restricted to sex-ratio lineages—see subsequently) (Stouthamer, Breeuwer, and Hurst 1999). The third class of association is mutualism, which is exemplified by the interaction between *Wolbachia* and their filarial nematode worm hosts. Unlike the parasitic Rickettsiaceae, if the *Wolbachia* infection of nematode worms is cured with tetracycline, the hosts themselves are harmed (Bandi et al. 1999). The phylogenies of the bacteria and their nematode hosts are congruent, suggesting that these associations have persisted for far greater periods of time than did the parasitisms (Bandi, Anderson, and Blaxter 1998).

The arthropod *Wolbachia* mostly fall into two phenotypic classes, which are intermingled on phylogenies of bacterial genes (Zhou, Rousset, and O'Neill 1998). In the first type are the sex-ratio distorters, in which the bacterium increases the production of daughters at the expense of sons. This benefits the bacteria, which are transmitted only through females, but is detrimental to the host, whose fitness is usually maximized by investing equally in sons and daughters (Cosmides and Tooby 1981). Selection, therefore, acts to promote host genes that prevent parasite action and transmission. In the second class of *Wolbachia* are those that induce CI. These strains of *Wolbachia* invade the host population by reducing the reproductive success of uninfected females, thereby increasing the frequency of the infected cytoplasm in the population. In diploid hosts they cause a

Abbreviation: CI, cytoplasmic incompatibility.

Key words: coevolution, *Cowdria*, *map1* gene, positive selection, Rickettsiaceae, *Wolbachia*, *wsp* gene.

Address for correspondence and reprints: Francis M. Jiggins, Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, United Kingdom.
E-mail: fmj1001@mole.bio.cam.ac.uk.

Mol. Biol. Evol. 19(8):1341–1349. 2002

reduction in the viability of eggs when infected males mate with females that do not carry the same strain of *Wolbachia*. This reduces the fitness of uninfected females in the population, thereby indirectly increasing the relative fitness of the infected females and leading to the spread of the infection through the population (Caspari and Watson 1959; Turelli and Hoffmann 1991). The extent of host-parasite conflict is more limited in these associations. Selection on both the bacterium and the female hosts (the transmitting sex) tends to act to increase both the transmission efficiency of the bacterium and the level of compatibility between infected males and females (Turelli 1994). Therefore, there may be little conflict between CI *Wolbachia* and their hosts.

Within the Rickettsiaceae, therefore, we would predict that the level of host-symbiont conflict would vary between the different types of symbiosis. In mutualistic interactions there is little conflict, and there is unlikely to be any selection on the host either to suppress the bacterial phenotype or to eliminate the infection. Therefore, comparing the major taxonomic groups, we expect host parasite arms races in the arthropod *Wolbachia* and *C. ruminantium* but not in the nematode *Wolbachia*. At a finer taxonomic scale, comparing the strains of arthropod *Wolbachia*, we expect a greater amount of host-parasite conflict in sex ratio–distorting strains than in those that induce CI.

These hypotheses can be tested by investigating the nature of selection acting on a bacterial gene that is likely to be involved in host-parasite interactions. We used the outer membrane proteins *map1* in *C. ruminantium* (Vanvliet et al. 1994) and its homologue *wsp* in *Wolbachia* (Braig et al. 1998). The *map1* gene is a member of a multigene family the members of which are closely related to genes found in rickettsial parasites of humans and canines (Sulsona, Mahan, and Barbet 1999). Phylogenetic analysis of these genes suggests that they have diverged after the split from *Wolbachia*, although this result relies on the uncertain rooting of the tree (Ohasi et al. 1998). Outer membrane proteins are likely to be involved in host-parasite interactions, and this has been confirmed in the case of *map1*, which is known to act as an antigen (Vanvliet et al. 1994; Perez et al. 1998). To investigate whether natural selection has favored diversification in these proteins along specific lineages of the bacterial phylogeny, or at specific amino acid sites, we used maximum likelihood analysis of a model of codon substitution to reconstruct the ratio of nonsynonymous to synonymous nucleotide substitutions (the d_N/d_S ratio) (Yang and Bielawski 2000). Neutral evolution will produce a d_N/d_S value of 1, whereas $d_N/d_S < 1$ indicates purifying selection, and $d_N/d_S > 1$ indicates positive selection. Studies of the *wsp* gene have shown that it evolves faster in arthropods than in nematodes (Bazzocchi et al. 2000*b*), but positive selection was not detected in the arthropod sequences (Zhou, Rousset, and O'Neill 1998). More recent analyses of a larger data set also found that in the majority of pairwise comparisons $d_N/d_S < 1$ (Schulenburg et al. 2000). However, the pairwise estimation of the average d_N/d_S ratio along the entire length of a protein is a highly conservative criterion for detecting positive selection.

There were three specific a priori hypotheses tested for these genes. First, the surface protein wsp in *C. ruminantium* and arthropod *Wolbachia* is positively selected, whereas the protein in nematode *Wolbachia* is not. Second, positive selection is stronger on sex ratio–distorting strains than on CI strains. Third, these differences are in regions of the protein that are likely to interact with the host.

Methods

Sequence Data

We used published sequences of the *wsp* gene from *Wolbachia* and the *map1* gene from *C. ruminantium*. The sequences were used to create four separate alignments (alignment done by eye) containing the nematode *Wolbachia* (10 sequences), the arthropod *Wolbachia* (52 sequences), the *C. ruminantium* (14 sequences), and a fourth alignment containing all the *Wolbachia* (62 sequences).

In the case of the arthropod *Wolbachia*, only sequences for which the bacterium's probable phenotype was known were included. In this alignment the codons were numbered according to the reference sequence from strain *wRi* (accession number AF020070). Because of the missing data, codons 1–36 and 223–229 were excluded from the alignment; of these, codons 1–24 are a signal sequence that is cleaved from the mature peptide (Braig et al. 1998). In addition, codons 59–65 and 204–213 were excluded from this alignment because they contained numerous insertions and deletions, making the alignment of homologous sites uncertain.

The alignment of the *C. ruminantium* sequences included all the amino acid sites in the mature peptide, and the codons are numbered according to the reference sequence Welgevonden (accession number U49843). By homology with the *Wolbachia* sequence, codons 1–25 in the alignment are the signal peptide. The codons of *Wolbachia* from filarial nematodes were numbered according to the reference sequence from *Brugia malayi* (accession number AJ252061). The beginning and the end of the full protein sequence were missing from the alignment (codons 1–54 and 203–240). Of these, codons 1–24 are the signal peptide. When both the arthropod and nematode sequences were aligned, it was necessary to exclude slightly different regions from those omitted in the previous alignments. This alignment, numbered according to reference sequence *wRi*, excluded codons 1–65 and 202–230.

The four alignments have been submitted to the EMBL alignment database under the accession numbers ALIGN_000198–ALIGN_000201. These alignments contain the accession numbers of the individual sequences used.

Tree Reconstruction

The phylogeny of the sequences in each of the four alignments was reconstructed by maximum likelihood using the program PAUP* v.4.0b8 (Swofford 1998).

First, the model of sequence evolution was selected for each alignment by comparing the likelihood scores using the program Modeltest v.3.04 (Posada and Crandall 1998). The models used were HKY85+G for all the *Wolbachia* alignments and GTR+I for the *C. ruminantium* alignment. For each alignment, model parameters were estimated from a maximum parsimony tree and then fixed in the maximum likelihood tree search using the nearest neighbor interchanges. These trees were then used for the estimation of d_N/d_S by maximum likelihood.

To assess the sensitivity of the results to the tree topology, a maximum parsimony tree was also generated, and all the branches that were supported by less than 50% of 100 bootstrap replicates were collapsed into multifurcations. These trees were used to replicate the results based on the maximum likelihood topologies.

The d_N/d_S Ratio Across Sites

The first hypothesis that was tested was that sites in the arthropod *Wolbachia* and *C. ruminantium* sequences are under positive selection, whereas positive selection is absent in the nematode *Wolbachia*. A maximum likelihood model of codon substitution was used to identify amino acid sites under positive selection (Nielsen and Yang 1998; Yang et al. 2000). The alignments of sequences from arthropod *Wolbachia*, nematode *Wolbachia*, and *C. ruminantium* were analyzed using the codeml program in the PAML package (v3.0a [Yang 1997]). The distribution of the d_N/d_S ratio (ω) across sites was estimated for the alignments of sequences from arthropod *Wolbachia*, nematode *Wolbachia*, and *C. ruminantium*. Four different models of codon substitution were used (Yang et al. 2000). M0 (one ratio) assumes that all sites have the same value of ω. M3 (discrete) assumes three different classes of sites with different ω ratios. M7 (beta) allows sites to have 10 different values of ω, calculated from the beta distribution with parameters p and q. The beta distribution is bounded between 0 and 1 and, thus, constitutes a null model for testing positive selection. M8 (beta + ω) is similar to M7 but with an additional ω category that can exceed 1.

The hypothesis that some amino acid sites are under positive selection can be tested by comparing two nested models with a likelihood ratio test, where these models differ in whether or not they allow some sites to have values of ω greater than 1. In this case M7 is compared with M8, and M3 with M0. The former comparison (M7 and M8) is the most stringent test of positive selection (Anisimova, Bielawski, and Yang 2001). The null distribution of the likelihood ratio test statistic ($2\Delta l$, where Δl is the difference between the log-likelihood scores of the two models) can be approximated using the χ^2 distribution with the degree of freedom (df) being the difference in the number of free parameters between the two models (df = 2 when comparing M7 and M8, and df = 4 when comparing M0 and M3). Analyses of simulated data sets indicate that this test is conservative (Anisimova, Bielawski, and Yang 2001).

When the likelihood ratio tests suggest the presence of sites under positive selection, an empirical Bayes method is used to calculate the posterior probabilities that each site falls into the ω classes (Yang et al. 2000). Sites with high probabilities of belonging to the class with $\omega > 1$ are likely to be under positive selection.

Prediction of Transmembrane Regions

The hypothesis that positive selection results from host-parasite interactions predicts that the positively selected regions of the protein will mostly be in the outer membrane region of the protein. To test this, the transmembrane regions of the protein were predicted from the sequence by using three different methods, DAS (Cserzo et al. 1997), MemSat (Jones, Taylor, and Thorton 1994), and TMPred (Hofmann and Stoffel 1993). To ensure the results, robust predictions were made using all three programs for four arthropod *Wolbachia* sequences and four *C. ruminantium* sequences. These sequences were chosen because they covered all, or nearly all, the mature protein.

The d_N/d_S Ratio Along Lineages

The hypothesis that the ω ratio differs between arthropod and nematode sequences was tested directly using a likelihood ratio test comparing models allowing different ω ratios along branches within the phylogeny (Yang 1998; Yang and Nielsen 1998). The ω ratio was estimated using the combined *Wolbachia* alignment either by forcing all branches to have the same ω or by allowing the nematode and arthropod clades to have different ωs. The likelihood scores were then compared using a likelihood ratio test, assuming that the test statistic follows a χ^2 distribution with df = 1.

The final hypothesis tested was that, within the arthropod *Wolbachia*, the selection on the *wsp* gene is stronger in sex ratio–distorting strains than in CI strains. The two phenotypes are scattered across the phylogeny of the gene. First, the phenotype of internal branches was reconstructed by maximum parsimony. The value of ω was then estimated either by forcing each branch to have the same value of ω or by allowing sex-ratio and CI branches to have two different values of ω. The hypothesis that ω is different between CI and sex-ratio strains was tested by comparing the likelihood scores of the one-ratio and the two-ratio models using a likelihood ratio test, with the test statistic compared with a χ^2 distribution with df = 1.

Results

Positive Selection in Parasitic and Mutualistic Rickettsiaceae

To illustrate the relationships of the different bacterial strains, the phylogenies of *map1* and *wsp* are shown in figures 1 and 2. In accordance with our predictions we detected positive selection in the arthropod *Wolbachia* and the *C. ruminantium* sequences but not in the nematode *Wolbachia*. The likelihood scores of the different codon substitution models and sites under se-

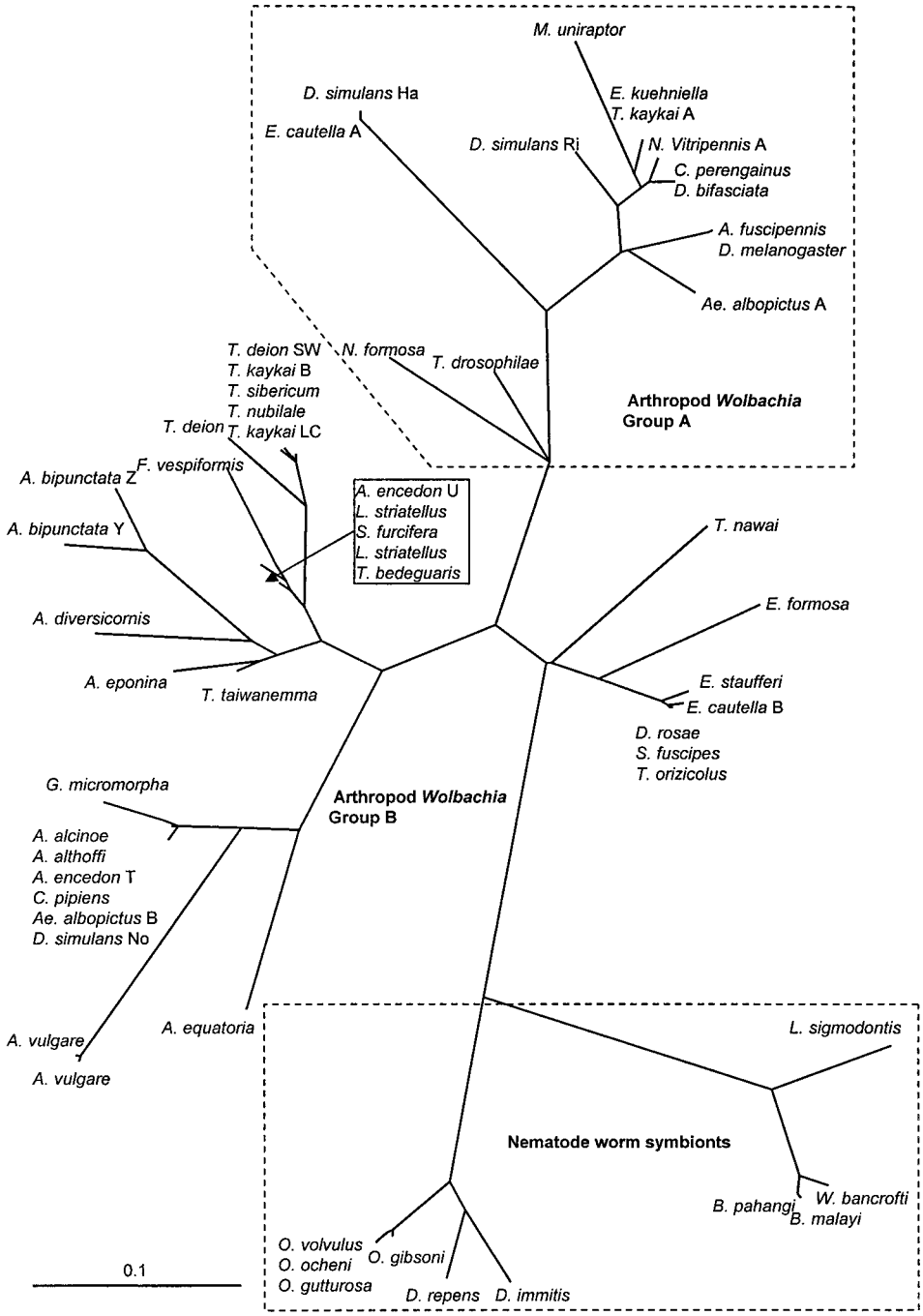

FIG. 1.—The maximum likelihood phylogeny of the *wsp* gene of *Wolbachia* bacteria from both arthropod and nematode hosts. The taxa are labeled with the name of the host organism or the bacterial strain, and the scale refers to the mean number of substitutions per nucleotide site.

lection are shown in tables 1–3. In the arthropod *Wolbachia*, likelihood ratio tests indicated that models that included positive selection significantly increased the likelihood scores when compared with models with no positive selection (M0 vs. M3: $2\Delta l = 565.32$, df = 4, $P < 0.001$; M7 vs. M8: $2\Delta l = 91.62$, df = 2, $P < 0.001$). Similarly, these tests supported the presence of positive selection in the *C. ruminantium* sequences (M0 vs. M3: $2\Delta l = 185.5$, df = 4, $P < 0.001$; M7 vs. M8: $2\Delta l = 7.7$, df = 2, $P < 0.05$). These results contrast with those from the nematode *Wolbachia*, in which no sites under positive selection were detected (table 2). The difference in the d_N/d_S ratio at homologous sites in the *wsp* gene from arthropod and nematode *Wolbachia* is shown graphically in figure 3.

We directly tested whether the value of ω averaged across sites was greater in the arthropod than in the nematode sequences by comparing the likelihood score of a model in which nematode and arthropod lineages had the same ω against a model in which ω differed between

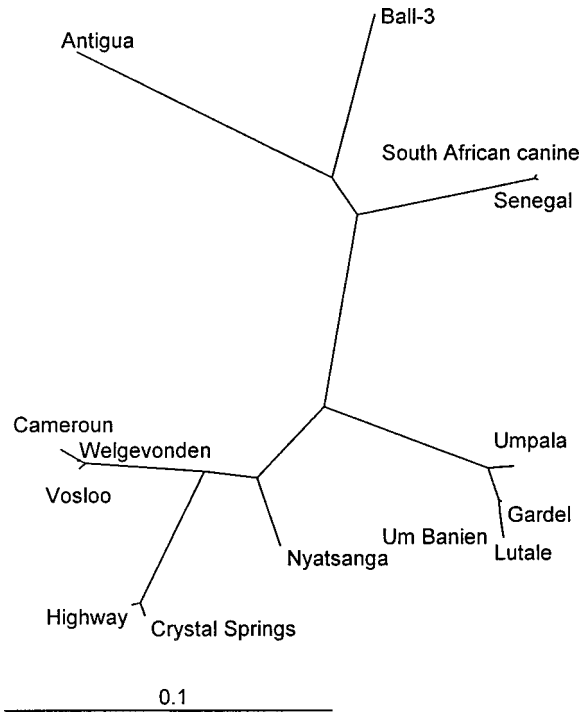

FIG. 2.—The maximum likelihood phylogeny of the *map1* gene from *Cowdria ruminantium*. The taxa are labeled with the names of the strains, and the scale refers to the mean number of substitutions per nucleotide site.

the two clades. The two-ratio model estimated ω to be greater in arthropods than in nematodes (table 4), and this model significantly increased the likelihood score relative to the one-ratio model ($2\Delta l = 15.86$, df = 1, $P < 0.001$).

Finally, we investigated the robustness and consistency of these results. First, we analyzed the sequences from the A and B groups (see fig. 1) of arthropod *Wolbachia* separately. In both cases positive selection was detected both by comparing M7 versus M8 and M0 versus M3 ($P < 0.001$ in all cases; data not shown). Furthermore, similar regions of the protein were under positive selection in both groups; the eight positively selected codons in the A group were also detected in the B group. A smaller number of positively selected sites was detected in the A group (8) compared with the B group (20) or with the combined data (20). This is probably because of the high estimate of ω in the A-group analysis (ω = 9), meaning that fewer sites were assigned to this category. Second, we investigated the sensitivity of our results to the tree topology by replicating the analysis using a maximum parsimony bootstrap consensus tree. The results under this tree topology were virtually identical to those obtained using the maximum likelihood tree. Other factors such as recombination or concerted evolution may further confound the phylogeny, and therefore we replicated the analysis for a third time using a star phylogeny. Again, positive selection was detected in the parasitic but not in the mutualistic data sets. Third, different regions of the protein were excluded from the analysis of the arthropod and nematode *wsp* sequences (fig. 1). We repeated the analysis including only homologous codons in both data sets. Again, this had little effect on the results. Finally, the nematode *Wolbachia wsp* data set had fewer taxa and a shorter tree length. Therefore, we repeated the analysis on four alignments of between seven and ten arthropod *Wolbachia* sequences with a tree length similar to that of the nematode data set. In all four replicates, positive selection was detected.

Prediction of Transmembrane Regions

Two transmembrane regions were consistently predicted for each of the four arthropod *wsp* sequences. In the reference sequence (*wRi*) these were inferred to be at codons 117–125 and 140–147 by DAS, at 111–127 and 138–150 by MemSat, and at 111–130 and 136–153 by TMPred. In the other three sequences examined, at least two of the three programs predicted these two transmembrane regions. Furthermore, these regions were similar to those predicted in an earlier study (Braig et al. 1998). The methods failed to produce a consistent prediction as to the orientation of the transmembrane regions. Consistent with our hypothesis, the transmembrane region is not subject to positive selection (fig. 1). The predicted transmembrane regions for the four *C. ruminantium* sequences used were not consistent across

Table 1
Positively Selected Sites, Log-Likelihood Scores, and Parameter Estimates for the *wsp* Gene from Arthopod *Wolbachia*

Model	Parameters in the ω Distribution	l	Positively Selected Codons
M0 (one ratio)...	$\omega = 0.34$	−4992.35	None
M3 (discrete)....	$\omega_1 = 0.06$, $p_1 = 0.53$ $\omega_2 = 0.46$, $p_2 = 0.36$ $\omega_3 = 2.40$, ($p_3 = 0.11$)	−4709.69	**51F, 58E, 66H, 97Q, 102D, 107A, 108F, 109N, 111D, 155K, 156D, 157A, 158V, 163S, 201K, 202D, 203T**
M7 (beta).....	$p = 0.32$, $q = 0.67$	−4757.59	Not allowed
M8 (beta + ω)...	$p = 0.58$, $q = 2.01$, $p_1 = 0.89$ $\omega = 2.46$ ($p_2 = 0.11$)	−4711.78	**51F,** 58E, **66H, 97Q, 102D, 107A, 108F,** 109N, **111D, 155K, 156D,** 157A, **158V, 163S, 201K, 202D, 203T**

NOTE.—Sites are identified by both the number and the amino acid in the reference sequence (*wRi*). The parameters p and q describe the shape of the beta distribution of ω, and p_1, p_2, and p_3 are the proportions of codons belonging to each category. Proportions that are not free parameters are in parentheses. Positively selected codons are those at which the posterior probability that the codon belongs to the positively selected class is $P > 0.95$, with those at which $P > 0.99$ shown in bold.

Table 2
Positively Selected Sites, Log-Likelihood Scores, and Parameter Estimates for the *wsp* Gene from Nematode *Wolbachia*[a]

Model	Parameters in the ω Distribution	*l*	Positively Selected Codons
M0 (one ratio)	$\omega = 0.14$	−1243.96	None
M3 (discrete)	$\omega_1 = 0.00$, $p_1 = 0.40$ $\omega_2 = 0.17$, $p_2 = 0.45$ $\omega_3 = 0.61$ ($p_3 = 0.15$)	−1234.27	None
M7 (beta)	$p = 0.37$, $q = 1.79$	−1234.48	Not allowed
M8 (beta + ω).	$p = 0.14$, $q = 0.81$ $p_1 = 0.67$ $\omega = 0.22$, ($p_2 = 0.33$)	−1234.38	None

[a] See note to table 1.

methods and sequences, and these results were, therefore, discarded.

Selection Pressure on Sex Ratio–Distorting and CI Strains

The second hypothesis tested was whether positive selection was stronger along sex-ratio *Wolbachia* lineages than along CI *Wolbachia* lineages. First, internal branches were assigned to either phenotype by maximum parsimony. The value of ω was then estimated using a two-ratio model in which ω differed between sex-ratio and CI branches. Contrary to our prediction, ω was estimated to be slightly higher in the CI branches ($\omega_{CI} = 0.37$ vs. $\omega_{SR} = 0.33$; table 5). When this model is compared with the one-ratio model, there is no evidence of any difference in the d_N/d_S ratio along CI and sex-ratio lineages ($2\Delta l = 0.32$, df = 1, not significant).

Discussion

We have tested two a priori hypotheses. The predictions of the first, that host-symbiont conflicts are stronger in the parasitic than in the mutualistic Rickettsiaceae, were supported. However, the predictions of the second, that conflict is greater in sex ratio–distorting strains than in CI strains, were not supported.

Parasites and Mutualists

The first hypothesis was that antagonistic coevolution between the host and the parasitic Rickettsiaceae should exert stronger selection on the symbionts when compared with mutualistic relationships. As predicted, the outer membrane protein studied was under positive selection in the parasites but not in the mutualists. Furthermore, sites predicted to be under selection fell outside the predicted transmembrane regions and, hence, possibly in regions of the protein exposed to the host. The importance of the *map1* gene in host-parasite interactions has been confirmed experimentally in *C. ruminantium*, where variation in the sequence of this protein has been shown to correlate with variable antigenic properties (Perez et al. 1998).

An important limitation of this analysis is that we have analyzed data only from a single gene in two parasitic and in one mutualistic bacterial clades. This clearly raises the possibility that some other factors have confounded our analysis, such as a change in protein function or changes in the ecology of the bacteria other than between parasitism and mutualism. Therefore, the conclusion that parasitism but not mutualism selects for accelerated change in protein sequence needs to be replicated both across additional genes and for further bacterial symbioses. The conclusion that the positive selection pressure is from the host immune system will need to be confirmed by functional studies because convincing support for this idea is available only for *C. ruminantium* (Perez et al. 1998). The possible interactions between the arthropod host and *Wolbachia* remain a matter for speculation. However, surface proteins may play a role in recognition by the host and also potentially as targets of the host in attempts to resist parasite proliferation or the entry of the bacteria into the egg.

Table 3
Positively Selected Sites, Log-Likelihood Scores, and Parameter Estimates for the *map*1 Gene from *Cowdria ruminantium*

Model	Parameters in the ω Distribution	*l*	Positively Selected Codons
M0 (one ratio)	$\omega = 0.11$	−2875.65	None
M3 (discrete).	$\omega_1 = 0.00$, $p_1 = 0.60$ $\omega_2 = 0.16$, $p_2 = 0.31$ $\omega_3 = 1.27$, ($p_3 = 0.09$)	−2782.88	**81S, 82E,** 84T, 85N, **152Q, 154S, 155A, 160T, 162A,** 258A, 262V, 263S
M7 (beta)	$p = 0.11$, $q = 0.67$	−2787.07	Not allowed
M8 (beta + ω)	$p = 0.22$, $q = 3.00$, $p_1 = 0.92$ $\omega = 1.40$ ($p_2 = 0.08$)	−2783.23	81S, 82E, 152Q, **154S,** 155A

NOTE.—Sites are identified by both the number and the amino acid in the reference sequence Welgevonden. See note to table 1.

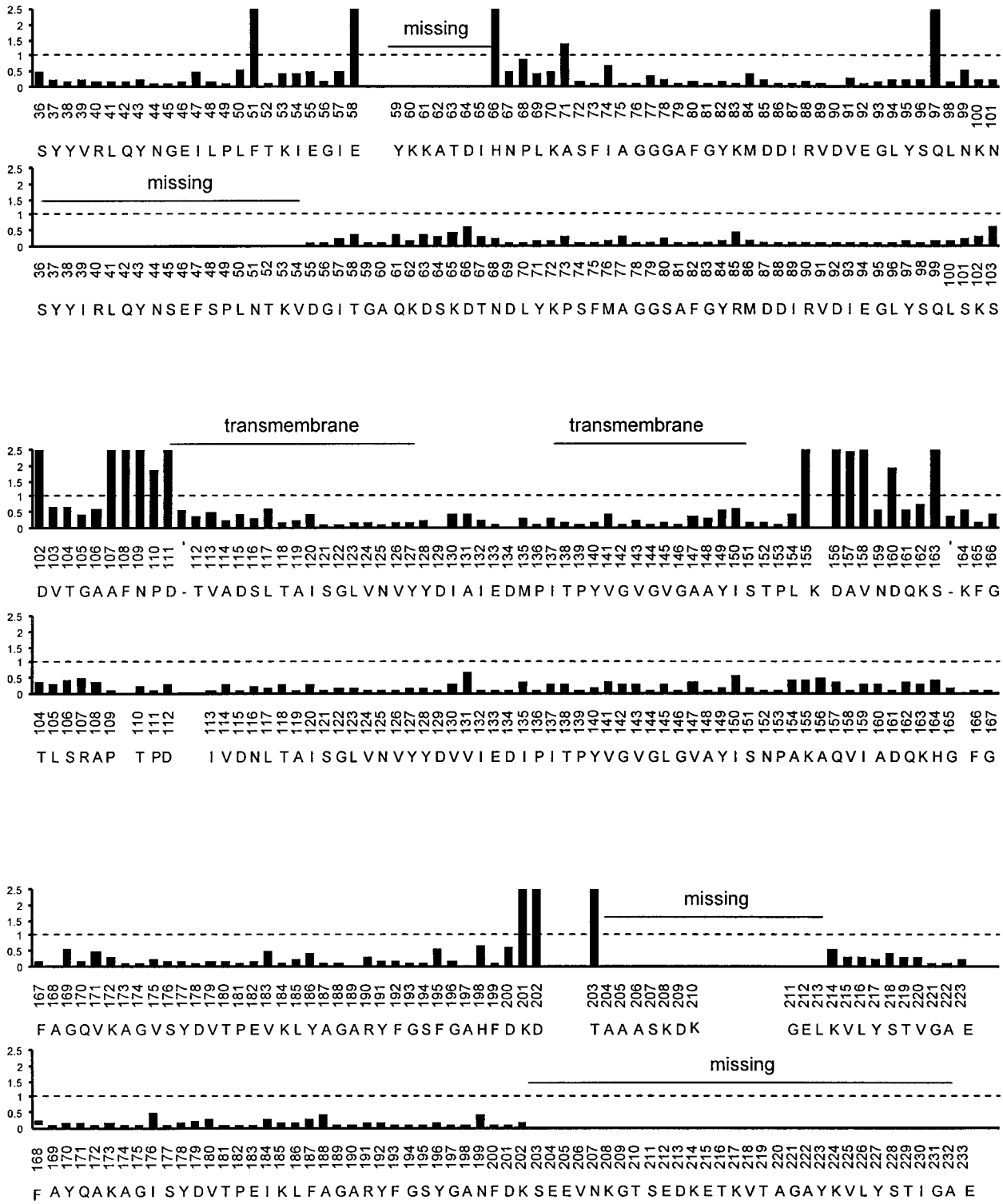

FIG. 3.—The predicted d_N/d_S ratio at homologous amino acid sites in the *wsp* gene of the parasitic arthropod *Wolbachia* (top) and the mutualistic nematode *Wolbachia* (bottom). This is calculated as the average of the ω ratios over the 11 classes from model M8, with the posterior probabilities used as weights. This is an approximation of the posterior mean of ω for the site. Homologous amino acid sites were aligned for the figure using the program SIM (Huang and Miller 1991). Also marked are the putative transmembrane regions of the arthropod protein (estimated by DAS) and the regions of the sequence that were missing from the analysis, both of which are described in the text.

It has been suggested that *map*1 could be a candidate for vaccine development (Perez et al. 1998). Our analysis has identified sites likely to be under positive selection. One simple explanation is that those sites are involved in antigenic variation in *map*1 and, thus, evolve at fast rates because of the selective pressure from the host immune response. Identification of those sites may thus provide insights into future efforts of vaccine de-

Table 4
Differences in the d_N/d_S Ratio of the *wsp* Gene in Nematode *Wolbachia* (mutualists) and Arthropod *Wolbachia* (parasites) Sequences

Model	Parameters	l
One ratio.....	$\omega = 0.27$	−4789.27
Two ratios....	$\omega_{nem} = 0.12$, $\omega_{arth} = 0.31$	−4781.34

NOTE.—In the one-ratio model all lineages have the same ω ratio. In the two-ratio model the nematode clades and arthropod clades have separate ω ratios (ω_{nem} and ω_{arth}, respectively).

Table 5
Differences in the d_N/d_S Ratio of the *wsp* Gene of Sex Ratio–Distorting and Cytoplasmic incompatibility (CI) *Wolbachia*

Model	Parameters	Likelihood
One ratio.....	$\omega = 0.34$	−4992.31
Two ratios....	$\omega_{SR} = 0.33$, $\omega_{CI} = 0.37$	−4992.15

NOTE.—In the one-ratio model all lineages have the same ω ratio. In the two-ratio model the sex ratio–distorting lineages and CI lineages have separate ω ratios (ω_{SR} and ω_{CI}, respectively).

velopment. Similarly, the positively selected sites in the *wsp* gene (the *map1* homologue) of arthropod *Wolbachia* are also likely to be important in host-parasite interactions. Our structural predictions located transmembrane regions of the protein, which are predicted to be under stabilizing or purifying selection in the maximum likelihood analysis. The structural prediction and the maximum likelihood comparison are thus consistent as we expect the transmembrane regions to be conserved.

The filarial nematode worms infected with mutualistic *Wolbachia* are themselves parasites of vertebrates. However, because the bacteria are within the cells of the worm, it is perhaps less likely (although not impossible) that the vertebrate immune system will exert any selection pressure on the outer membrane proteins of *Wolbachia*. Despite this, cats infected with filarial nematodes produced antibodies to the *wsp* protein (Bazzocchi et al. 2000*a*). There may be an immune response to *Wolbachia* outside the worm that are probably released from dead nematodes. Indeed, treatment of filarial nematode infections is known to cause an inflammatory response in mice through the release of *Wolbachia* from dead worms (Taylor, Cross, and Bilo 2000). The absence of positive selection on the *wsp* gene of these bacteria is consistent with the hypothesis that this vertebrate immune response does not result in the elimination of the nematode worms. Testing of this hypothesis is timely.

Sex-Ratio Distortion and CI

The second hypothesis was that, within the arthropod *Wolbachia*, the selection on sex-ratio strains would be greater than on strains that induce CI. This is expected because, all else being equal, infection with a sex-ratio distorter is deleterious. In contrast, there may be a cost for being an uninfected female in a population infected with CI *Wolbachia* because such females cannot mate successfully with infected males. However, this hypothesis was not supported by our analysis because we could not detect any difference in the d_N/d_S ratio in sex-ratio and CI strains. A possible factor limiting the power of our test is the observation that strains may switch fairly frequently between sex-ratio distortion and CI (F. Jiggins, unpublished data). If this is the case, then our assignment of internal and even terminal branches of the tree to a particular phenotype may have been inaccurate.

However, if selection on CI and sex-ratio strains is similar, then positive selection on the arthropod *wsp* sequences may result from aspects of the parasitic lifestyle that are held in common by both the phenotypic classes. In particular, parasitism is associated with frequent shifts between host species, whereas mutualism leads to long-term stable relationships involving cocladogenesis of host and bacterium (Bandi, Anderson, and Blaxter 1998; Zhou, Rousset, and O'Neill 1998). Therefore, continual adaptation to novel hosts may, in part, account for the positive selection on *wsp* in arthropod *Wolbachia*. It is noteworthy that the host range of *C. ruminantium* is probably intermediate between the two *Wolbachia* data sets.

The validity of all these conclusions depends on our ability to accurately estimate the d_N/d_S ratio. One point of concern is that the phylogeny of the *wsp* gene is known not to be very robust (Schulenburg et al. 2000), and factors such as recombination within the gene may confound the reconstruction of the tree (Jiggins et al. 2001; Werren and Bartos 2001). However, this problem does not appear serious because the method used is known to be insensitive to the tree topology (Yang et al. 2000; Swanson et al. 2001). This was confirmed for our data set, where different trees all produced similar results.

Acknowledgments

We thank two anonymous reviewers and the associate editor for their comments on this manuscript. F.M.J. is funded by a Research Fellowship from Emmanuel College, Cambridge, and a Natural Environment Research Council small grant. G.D.D.H. is supported by a Biotechnology and Biological Sciences Research Council D. Phillips Fellowship. Z.Y. is supported by a Biotechnology and Biological Sciences Research Council grant and an EU Human Frontier Science Programme grant.

LITERATURE CITED

ANISIMOVA, M., J. P. BIELAWSKI, and Z. YANG. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol. Biol. Evol. **18**:1585–1592.

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

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.

BAZZOCCHI, C., F. CECILIANI, J. W. MCCALL, I. RICCI, C. GENCHI, and C. BANDI. 2000*a*. Antigenic role of the endosymbionts of filarial nematodes: IgG response against the Wolbachia surface protein in cats infected with Dirofilaria immitis. Proc. R. Soc. Lond. Ser. B Biol. Sci. **267**:2511–2516.

BAZZOCCHI, C., W. JAMNONGLUK, S. L. O'NEILL, T. J. C. ANDERSON, C. GENCHI, and C. BANDI. 2000*b*. wsp Gene sequences from the Wolbachia of filarial nematodes. Curr. Microbiol. **41**:96–100.

BRAIG, H. R., W. ZHOU, S. L. DOBSON, and S. L. O'NEILL. 1998. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. J. Bacteriol. **180**:2373–2378.

CASPARI, E., and G. WATSON. 1959. On the evolutionary importance of cytoplasmic sterility in mosquitoes. Evolution **13**:568–570.

COSMIDES, L. M., and J. TOOBY. 1981. Cytoplasmic inheritance and intragenomic conflict. J. Theor. Biol. **89**:83–129.

CSERZO, M., E. WALLIN, I. SIMON, G. VONHEIJNE, and A. ELOFSSON. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng. **10**:673–676.

EHRLICH, P. R., and P. H. RAVEN. 1964. Butterflies and plants: a study in coevolution. Evolution **18**:586–608.

HOFMANN, K., and W. STOFFEL. 1993. TMbase—a database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler **374**:166.

HUANG, X. Q., and W. MILLER. 1991. A time-efficient, linear-space local similarity algorithm. Adv. Appl. Math. **12**:337–357.

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

JONES, D. T., W. R. TAYLOR, and J. M. THORTON. 1994. A model recognition approach to the prediction of all-helical membrane-protein structure and topology. Biochemistry **33**: 3038–3049.

NIELSEN, R., and Z. YANG. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics **148**:929–936.

OHASI, N., A. UNVER, N. ZHI, and Y. RIKIHISA. 1998. Cloning and characterization of multigenes encoding the immunodominant 30-kilodalton major outer membrane proteins of *Ehlichia canis* and application of the recombinant protein for serodiagnosis. J. Clin. Microbiol. **36**:2671–2680.

PEREZ, J. M., D. MARTINEZ, C. SHEIKBOUDOU, F. JONGEJAN, and A. BENSAID. 1998. Characterization of variable immunodominant antigens of Cowdria ruminantium by ELISA and immunoblots. Parasite Immunol. **20**:613–622.

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

SCHULENBURG, J. H. G. V. D., 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 phylogentic utility of *WolbachiaftsZ* and *wsp* gene sequences with special reference to the origin of male-killing. Mol. Biol. Evol. **17**: 584–600.

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

SULSONA, C. R., S. M. MAHAN, and A. F. BARBET. 1999. The map1 gene of Cowdria ruminantium is a member of a multigene family containing both conserved and variable genes. Biochem. Biophys. Res. Commun. **257**:300–305.

SWANSON, W. J., Z. YANG, M. F. WOLFNER, and C. F. AQUADRO. 2001. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. Proc. Natl. Acad. Sci. USA **98**:2509–2514.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, Mass.

TAYLOR, M. J., H. F. CROSS, and K. BILO. 2000. Inflammatory responses induced by the filarial nematode Brugia malayi are mediated by lipopolysaccharide-like activity from endosymbiotic Wolbachia bacteria. J. Exp. Med. **191**:1429–1435.

TURELLI, M. 1994. Evolution of incompatibility-inducing microbes and their hosts. Evolution **48**:1500–1513.

TURELLI, M., and A. A. HOFFMANN. 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature **353**:440–442.

VANVLIET, A. H. M., F. JONGEJAN, M. VANKLEEF, and B. A. M. VANDERZEIJST. 1994. Molecular-cloning, sequence-analysis, and expression of the gene encoding the immunodominant 32-kilodalton protein of *Cowdria-ruminantium*. Infect. Immun. **62**:1451–1456.

WERREN, J. H., and J. D. BARTOS. 2001. Recombination in Wolbachia. Curr. Biol. **11**:431–435.

YANG, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13**: 555–556.

———. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. **15**:568–573.

YANG, Z. H., and J. P. BIELAWSKI. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. **15**: 496–503.

YANG, Z., and R. NIELSEN. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J. Mol. Evol. **46**:409–418.

YANG, Z. H., R. NIELSEN, N. GOLDMAN, and A. M. K. PEDERSEN. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics **155**:431–449.

ZHOU, W. F., F. ROUSSET, and S. O'NEILL. 1998. Phylogeny and PCR based classification of *Wolbachia* strains using *wsp* gene sequences. Proc. R. Soc. Lond. B **265**:509–515.

PEKKA PAMILO, reviewing editor

Accepted April 12, 2002