Copyright © 2005 by the Genetics Society of America
DOI: 10.1534/genetics.105.046342

An Ancient Mitochondrial Polymorphism in *Adalia bipunctata* Linked to a Sex-Ratio-Distorting Bacterium

Francis M. Jiggins[1] and Matthew C. Tinsley

Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received June 1, 2005
Accepted for publication July 22, 2005

ABSTRACT

Sex-ratio-distorting microbes are common parasites of arthropods. Although the reasons they have invaded and spread though populations are well understood, their subsequent dynamics within those populations are virtually unknown. We have found that different strains of a male-killing Rickettsia bacterium infecting the beetle *Adalia bipunctata* are associated with distinct mitochondrial haplotypes, which is expected as both the mitochondria and the bacteria are maternally transmitted. These mitochondrial haplotypes shared a common ancestor >2 million years ago, and their overall diversity is significantly greater than expected under neutrality from comparisons with a nuclear gene. Furthermore, a variety of statistical tests show strong deviations from neutrality in mitochondrial but not in nuclear genes. We therefore conclude that natural selection is probably maintaining a polymorphism of different Rickettsia strains in this species. Despite the age of the different mitochondrial haplotypes, there is very little genetic diversity within them. Furthermore, there is considerable variation in mitochondrial haplotype and bacterial strain frequency between populations, despite it being thought that this species has fairly low levels of population structure. We conclude that the fitness of these male killers may be negatively frequency dependent or different strains may be favored in different populations. These hypotheses await experimental confirmation.

AN evolutionary conflict exists between nuclear and cytoplasmic genes over the optimal ratio of sons to daughters. Cytoplasmic genes are maternally transmitted and therefore favor a female-biased sex ratio, while autosomal genes favor a 1:1 sex ratio (Cosmides and Tooby 1981). As predicted, a diversity of cytoplasmic elements that distort the sex ratio toward females have invaded animal and plant populations (Laser and Lersten 1972; Hurst 1993). In response, nuclear genes have evolved defenses against these sex-ratio distorters, by blocking their transmission or action (Cavalcanti *et al.* 1957; Malogolowkin 1958; Frank 1998).

The aim of this study was to explore the dynamics of sex-ratio-distorting elements within populations. One important factor that may influence the dynamics of sex-ratio distorters is antagonistic coevolution with nuclear genes. For example, theoretical models have shown that if there are specific gene-for-gene-type interactions between sex-ratio distorters and their hosts, then both the cytoplasmic and nuclear genes may be highly polymorphic (Frank 1989; Randerson *et al.* 2000). Furthermore, instead of reaching a stable equilibrium the two sets of genes may cycle in frequency through time (Charlesworth 1981; Gouyon *et al.* 1991). Alternatively, coevolution may also result in an arms race in which host and parasite genotypes are continually fixed, resulting in low levels of genetic variation in the population. However, evolutionary constraints may mean that hosts do not mount any defenses and sex-ratio distorters may simply go to equilibrium. Other factors, such as adaptation of sex-ratio distorters to local conditions (host or environment), may also determine the distribution and diversity of these elements.

Although these processes have received much attention in plants, they are seldom studied within animal populations. For this reason, we have examined the selection pressures acting on a male-killing bacterium found in the two-spot ladybird (ladybug) beetle, *Adalia bipunctata* (Werren *et al.* 1994). Male killers are intracellular, maternally transmitted, and normally kill the sons of infected females before they hatch from the egg. They can spread when the death of sons benefits the surviving daughters, as it is daughters that transmit the infection to the next generation (Hurst and Majerus 1993). This may occur if females cannibalize their dead brothers or are prevented from inbreeding or competing with them (Werren 1987; Hurst and Majerus 1993). In most species, imperfect vertical transmission of the bacterium occurs, resulting in the production of uninfected sons and daughters. The causes of imperfect transmission are mostly unknown, but it is thought to explain the coexistence of infected and uninfected females within populations (Hurst 1991).

In this study we have used patterns of mitochondrial DNA (mtDNA) polymorphism to infer the selection pressures acting on the cytoplasmic genome (*i.e.*, mtDNA

[1]*Corresponding author:* Institute of Evolutionary Biology, University of Edinburgh, Ashworth Lab, King's Bldgs., W. Mains Rd., Edinburgh EH9 3JT, Scotland. E-mail: francis.jiggins@ed.ac.uk

plus male killer). This is possible because both the bacteria and mtDNA are maternally transmitted and therefore remain in linkage disequilibrium. The cotransmission of the two genomes is known both from experimental studies of bacterial transmission and from the observation that they are in complete linkage disequilibrium (Hurst *et al.* 1992; von der Schulenburg *et al.* 2002; this article). Although horizontal transmission of male killers has occurred during the establishment of new infections, it must be sufficiently rare to maintain linkage disequilibrium with mtDNA (Hurst and Jiggins 2005). Therefore, selection acting on the bacterium will be reflected in the patterns of mtDNA polymorphism.

We are able to reject the null hypothesis that the cytoplasmic genes have evolved neutrally and show that multiple mitochondrial haplotypes have been maintained in the host population for long periods of evolutionary time. Furthermore, there is considerable spatial variation in the frequency of these haplotypes. We argue that these patterns are probably the result of natural selection maintaining variation in the male-killer population. This suggests that the fitness of male-killer strains may be negatively frequency dependent or that different strains are favored in different populations with a polymorphism resulting from selection/migration balance.

METHODS

Linkage disequilibrium between bacterial and mitochondrial haplotypes: It has been previously shown that male killers and mtDNA are in linkage disequilibrium within *A. bipunctata* populations (von der Schulenburg *et al.* 2002). However, that study did not differentiate different Rickettsia genotypes. Therefore we used PCR-RFLPs to genotype the mtDNA and bacteria from a sample of 137 Rickettsia-infected beetles collected from Ljusdal in Sweden. A variable region of the bacterial *ompB* gene was amplified with the primers *ompB*-forward (5′-GGTGATGCTGCTATTGCTGA) and *ompB*-reverse (5′-CTACGTTACCAGGGCCAGAA). The PCR product was then digested with the restriction enzyme *Pst*I. We also sequenced the *ompB* PCR product from 14 samples and submitted the sequences to GenBank under accession nos. AJ699163–AJ699164. The primers and enzymes used to genotype the mtDNA are described below.

Male-killer infection: We screened large population samples of female beetles for infection with the male killer. These were collected as adults or pupae in the summer of the years in question, sexed, and the males discarded. Collections were made from two sites in England (Cambridge and Box Hill in Surrey near London) and eight sites in Scandinavia along a 1400-km-long transect. Some additional samples were included from elsewhere in Europe. DNA was extracted from the specimens using a method modified from Walsh *et al.* (1991). Half of the abdomen was digested for 1 hr at 56° with 5% w/v Chelex 100 ion exchange resin (Bio-Rad, Hercules, CA) in 200 μl 33 mm dithiothreitol with 20 μg proteinase K, boiled, and the supernatant was used directly as the PCR template.

The samples were tested for infection with the Rickettsia bacteria using a diagnostic PCR reaction. The primers RSSUf and RSSUr (von der Schulenburg *et al.* 2001) amplify 380 bp of the 16S rDNA gene from Rickettsia bacteria. These were mixed with the primers C1-J-1751 (Simon *et al.* 1994) and C1-N-2618 (von der Schulenburg *et al.* 2002), which amplify part of the mitochondrial COI gene and act as a check that the DNA extraction and PCR reaction were successful. If the COI gene did not amplify, the sample was discarded. To check the results, a subset of the samples was also screened with two other diagnostic primer pairs that amplify the rickettsial 17-kDa surface antigen gene and citrate synthase gene, respectively (Williams *et al.* 1992; Davis *et al.* 1998).

Mitochondrial haplotypes: The mitochondrial haplotype of individuals in the population samples was determined by RFLP analysis. von der Schulenburg *et al.* (2002) described three mtDNA haplotypes associated with the Rickettsia male killer, which they named 9I, 7A, and 10J. We refer to these as haplotypes A, B, and C, respectively. An aliquot of the multiplex PCR reaction described above (RSSUf/RSSUr + C1-J-1751/C1-N-2618) was digested with a mixture of the restriction enzymes *Hae*III and *Rsa*I. The C haplotype produced bands of 51, 109, and 753 bp; the B haplotype bands were 51, 109, 232, and 521 bp; and the A haplotype bands were 51, 109, 365, and 388 bp. We checked that this test correctly identified the haplotypes in the United Kingdom (UK) population by sequencing 600 bp of the COI gene from seven of the A haplotypes from infected females, two of the B haplotypes from infected females, and four of the C haplotypes from uninfected females. In all cases the sequence was identical to that published in von der Schulenburg *et al.* (2002). It should be noted that this species is also host to male-killing Wolbachia and Spiroplasma in continental Europe (Hurst *et al.* 1999a,b), and the mtDNA sequences associated with these infections mostly have the same RFLP pattern (but differ by several nucleotide substitutions) as the type B mtDNA (von der Schulenburg *et al.* 2002). These additional bacteria are not known to occur in the English populations.

Differences in the frequency of the haplotypes in different populations were analyzed in contingency tables. Their significance was assessed by generating 100,000 random contingency tables with the same marginal values using a Monte Carlo procedure and taking the proportion with more extreme deviations as the probability (Lewontin and Felsenstein 1965).

To investigate the mtDNA variability associated with one bacterial strain, we also sequenced two longer

stretches of mtDNA from nine infected females with the A haplotype. The first of these covered most of the COI and COII genes together with the intervening tRNA genes. This was amplified in three sections using the primer pairs TY-J-1462/TL2-N-3014, C1-J-1951/TL2-N-3014, and C1-J-2630/TK-N-3795 (Simon *et al.* 1994; von der Schulenburg *et al.* 2002; Jiggins 2003). The second region covered most of the ND5 gene and some flanking tRNA genes. This was amplified with the primers N3-J-5945/N5-N-6777 (5′-CAGAATTAATAGGAGGACA GGG), N5-J-Adalia (5′-CCTTGATATCTTCAATATCAT GCTCT)/N5-N-Adalia (5′-CTTTTAGTTCATGCTTTG TTTAAGGC), and N5-J-7183/N5-N-7789. The PCR products were then sequenced using the same primers. The C1-J-1951/TL2-N-3014 was also sequenced from 10 A haplotype uninfected females.

One process that could introduce highly divergent mtDNA sequences into a population and confound our analyses is hybrid introgression from a related species. To guard against this possibility we also sequenced mtDNA from *A. decempunctata*, which is the sibling species of *A. bipunctata* and is also infected with a male-killing Rickettsia (Lus 1947; Iablokoff-Khnzorian 1982; von der Schulenburg *et al.* 2001). Mitochondrial sequences from the two species are known to form two monophyletic groups (von der Schulenburg *et al.* 2002). We confirmed this by sequencing the same set of mitochondrial genes from a Rickettsia-infected female *A. decempunctata* and aligning them with sequences from *A. bipunctata* and the outgroup *Coccinella septempunctata*. After reconstructing the phylogeny, there was 100% bootstrap support for the *A. bipunctata* sequences forming a monophyletic group. Therefore, the *A. bipunctata* sequences have not introgressed from *A. decempunctata*.

Comparing nuclear and mitochondrial diversity: To distinguish the effects of selection and demography, we compared patterns of polymorphism in the nuclear gene *g6pd* and mtDNA. The *g6pd* gene is X-linked (F. Jiggins, unpublished results). It encodes the enzyme glucose-6-phosphate dehydrogenase, which is involved in the oxidative stage of the pentose phosphate pathway. We discuss below how we correct for differences in male and female migration rates that could confound this comparison. We used published mtDNA sequences from Rickettsia-infected females that covered 610 bp of the COI gene and 563 bp of ND5 (von der Schulenburg *et al.* 2002). These 13 sequences are the only data that were not collected during the course of this study, and none of the analyses that we report have been described elsewhere. The nuclear gene *g6pd* was then sequenced from wild or F_1 specimens, collected wherever possible from the same locations as those in von der Schulenburg *et al.* (2002). The populations and number of alleles sequenced are shown in Figures 1 and 2.

The *g6pd* gene was amplified with the primers *g6pd-Adalia*-1F (5′-GGCAGATCATCTTGCTTCCTTATTCA) and *g6pd-Adalia*-3R (5′-GGNCCNCKNSWNCCRTATCT AAAAG), using *Taq* DNA polymerase and 30 thermal cycles. Each PCR product was ligated into a T-tailed vector and transformed into competent *Escherichia coli* strain JM109. The plasmids were purified from single clones, and the inserts were sequenced using primers designed from the flanking vector sequence and two internal primers. Copying errors during PCR were corrected by sequencing multiple clones from each specimen. Where we aimed to sequence only one allele, four clones were used, and when both alleles were required eight clones were used. Each clone was derived from a separate PCR reaction and transformation to avoid the potential duplication of errors either by replication of the bacteria before they are plated out or during the PCR reaction. The sequences have GenBank accession nos. AJ585242–AJ585257. A neighbor-joining tree was reconstructed from both the mtDNA and *g6pd* data sets on the basis of Jukes-Cantor (Jukes and Cantor 1969) distances and using *A. decempunctata* as an outgroup.

Data analysis: The statistical significance of a number of population genetic tests was obtained by coalescent simulation. The first approach we used was to generate random genealogies onto which were placed S mutations. We used this fixed-S method rather than simulations conditional on θ as S is a directly observed parameter, while θ must be estimated from the data. The fixed-S method has been criticized as the actual distribution of the statistic is dependent on both θ and S, and therefore the approach is invalid if θ is very different from that expected for a given S (Markovtsova *et al.* 2001). However, this study measured type I error rates for arbitrarily chosen combinations of S and θ, rather than values of θ within the confidence interval of θ given the observed S (Depaulis *et al.* 2001). Several approaches have subsequently supported the use of fixed-S simulations. First, Wall and Hudson (2001) found type I errors close to 5% using the fixed-S method applied to simulated data conditioned on θ. Second, Depaulis *et al.* (2001) found acceptable type I errors within the 95% confidence limits of θ. Furthermore, when the estimates of θ were weighted by their posterior probabilities, the fixed-S approach was conservative when applied to haplotype number or diversity tests (Depaulis *et al.* 2001). In conclusion, the fixed-S method is inaccurate only for values of S that are very improbable for a given value of θ (Depaulis *et al.* 2001). This is not a problem as the fact that the observed S is unexpected given θ is in itself a reason to reject the neutral null model (Depaulis *et al.* 2001; Wall and Hudson 2001).

Although the fixed-S approach has been shown to be robust, we also took another highly conservative approach. First, we calculated the 95% confidence limits of the Watterson (1975) estimator of θ, using the method of Kreitman and Hudson (1991). We then ran coalescent simulations conditioned on both the upper

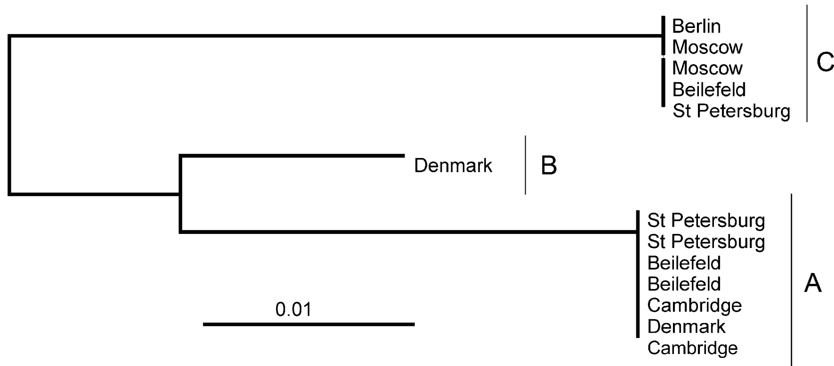

Figure 1.—Neighbor-joining tree of mitochondrial sequences from Rickettsia-infected females rooted with a sequence from *Adalia decempunctata* (not shown).

and lower confidence limits of θ. These simulations used the classical approach of generating random genealogies and then randomly placing mutations at the rate of θ/2 along each branch.

RESULTS

Linkage disequilibrium between bacterial and mitochondrial haplotypes: Within the sample of 137 Rickettsia-infected beetles from Ljusdal there is a perfect association between the bacterial and mitochondrial genotypes. The *ompB* PCR product from 14 beetles was cut by the enzyme *Pst*I, and all these individuals had the A or C mitotypes. The remaining 123 beetles had the B mitotypes, and the *ompB* gene did not have the *Pst*I cutting site. We then sequenced part of the *ompB* gene from 14 of the beetles. Seven beetles with the A mitotype had identical *ompB* sequences, as did 7 beetles with the B mitotype. However, there was considerable divergence between strains associated with different mtDNA haplotypes. In total, 26 of 441 bp differed between the *ompB* sequences associated with the A and B mtDNA.

Patterns of mitochondrial DNA polymorphism: The 13 mitochondrial sequences from a random sample of Rickettsia-infected beetles contained 78 segregating sites in 1172 bp of sequence. The per site nucleotide diversity of these sequences was very high, estimated as $\pi = 0.032$ and $\theta_W = 0.021$. The 95% confidence interval θ_W, estimated using the method of Kreitman and Hudson (1991), was 0.010–0.052. The genealogy of these sequences contains three clades, each separated by long internal branches (Figure 1). The sequences within these clades are identical. The genetic distance, estimated by the Jukes-Cantor (Jukes and Cantor 1969) method, is 0.06 between A and C and 0.05 between B and C. Assuming a pairwise divergence of 2.3% per million years (Brower 1994), C diverged from B and A 2.1–2.8 million years ago. The A and B clades are more closely related and are separated by a genetic distance of 0.03 or 1.4 million years.

There is extensive variation in the frequency of the three Rickettsia-infected cytotypes in different populations (Table 1; Monte Carlo simulation, $P < 0.000001$). These differences are particularly clear when comparing the large samples collected from different countries (Table 1). To establish the scale at which this variation occurs, we compared smaller samples from the same countries. In Cambridge, just 2 of 76 beetles were haplotype B while, at Box Hill just 115 km away, two of four infected females were haplotype B (Monte Carlo simulation, Cambridge and Box Hill, $P = 0.01$). A similar pattern is apparent in Scandinavia. Here the B haplotype predominates rather than the A haplotype found in Cambridge (Table 1). As was the case in England, there is significant heterogeneity between the different Swedish populations (Monte Carlo simulation, $P < 0.05$). Small samples from several other countries are shown in Table 1, and these also show marked variation in haplotype frequency.

To test whether there is also temporal variation in haplotype frequency, we compared the frequency of the two mtDNA haplotypes found in Cambridge in 1995, 2002, and 2003. The changes in frequency between years were very small, with type B being absent in the large 1995 sample but present in single specimens from the smaller 2002 and 2003 samples (Table 1). These changes were marginally significant (Monte Carlo simulation, $P < 0.05$), suggesting that frequencies may change through time. However, more data are needed before this result could be considered reliable. Larger sample sizes are available for the uninfected females, and here the changes in haplotype frequency over the 3 years had greater statistical support (Table 1; Monte Carlo simulation, $P = 0.007$). As discussed below, the haplotype frequency in uninfected females is related to that in infected females. It is therefore possible that the change in uninfected females reflects a change in infection frequency through time.

Evidence for selection maintaining variation: The shape of the mtDNA genealogy, with long internal branches and relatively short terminal branches, could result from selection maintaining cytoplasmic diversity or population structure (Takahata and Nei 1990; Vekemans and Slatkin 1994). It is possible to test for these effects because they will alter the frequency distribution of alleles within the population compared to the neutral expectation. The value of Tajima's (1989) D for the mtDNA data set is significantly >0 ($D = 2.22$,

TABLE 1

Mitochondrial haplotype frequencies in female *Adalia bipunctata*

				mtDNA haplotype		
Country	Site	Year	Infection	A	B	C
UK	Cambridge	1995	Rickettsia	48	0	0
UK	Cambridge	2002	Rickettsia	4	1	0
UK	Cambridge	2003	Rickettsia	22	1	0
UK	Box Hill	2002	Rickettsia	2	2	0
UK	Cambridge	1995	Uninfected	274	136	1
UK	Cambridge	2002	Uninfected	27	10	1
UK	Cambridge	2003	Uninfected	91	79	1
UK	Box Hill	2002	Uninfected	31	44	2
Sweden	Malmo	2001	Rickettsia	2	0	0
Sweden	Nassjo	2001	Rickettsia	0	8	0
Sweden	Stockholm	2000–2001	Rickettsia	0	1	0
Sweden	Gavle	2001	Rickettsia	0	7	0
Sweden	Ljusdal	2001	Rickettsia	13	123	1
Sweden	Ostersund	2000–2001	Rickettsia	0	15	0
Sweden	Vilhelmina	2001	Rickettsia	1	17	0
Norway	Narvik	2000	Rickettsia	0	3	0
Russia	Moscow		Rickettsia	0	0	2
Russia	St Petersburg		Rickettsia	2	0	1
Germany	Berlin		Rickettsia	0	0	1
Germany	Beilefeld		Rickettsia	2	0	0
Denmark	Copenhagen		Rickettsia	1	1	0

Table 2). This indicates that there is an excess of polymorphic sites at intermediate frequency within the population, as expected under balancing selection or in subdivided populations.

Fu and Li's (1993) D- and F-statistics test whether a genealogy fits the neutral model. These tests compare the number of mutations on internal branches with those on terminal branches. Relative to a neutral model of evolution, directional selection would result in an excess of mutations on terminal branches, while balancing selection may result in an excess of internal mutations. We used the mtDNA sequence from *A. decempunctata* as an outgroup to determine the ancestral state of the polymorphisms. Both of these tests indicated an excess of mutations on internal branches ($D = 1.54$, $F = 2.01$, Table 2).

TABLE 2

The probabilities of the observed values of test statistics under the neutral model

Test statistic	$S = 78^a$	$\theta = 12.2^b$	$\theta = 25.1^c$	$\theta = 60.9^b$
Tajima's $D = 2.22$	0.006	0.007	0.005	0.004
Fu and Li's $D = 1.54$	0.04	0.04	0.03	0.03
Fu and Li's $F = 2.01$	0.02	0.02	0.02	0.01
$H_d = 0.603$	0.00002	0.003	0.00008	$<10^{-5}$
$K = 3$	$<10^{-5}$	0.0003	$<10^{-5}$	$<10^{-5}$

All coalescent simulations were run 10^4 times and based on 13 sequences.

[a] Coalescent simulations using the fixed-S approach.

[b] Simulations conditioned on the upper and lower 95% confidence limits of θ estimated from S.

[c] Simulations conditioned on the Watterson (1975) estimate of θ.

Balancing selection can also result in the polymorphic sites in the sample being partitioned into fewer haplotypes than would be the case if they were evolving neutrally. This is the case for our data set, as both the haplotype diversity (H_d) (Nei 1987) and the number of haplotypes (K) (Nei 1987; Depaulis and Veuille 1998) were lower than the neutral expectation ($H_d = 0.603$, $K = 3$, Table 2).

Demographic processes such as bottlenecks and population structure can alter the frequency distribution of alleles in ways similar to selection. To distinguish demography from selection, we sequenced the nuclear gene *g6pd*. The genealogy of *g6pd* (Figure 2) contrasts with the mtDNA genealogy in that the internal branches are shorter relative to the terminal branches. Furthermore, none of the statistics above differed significantly from neutrality (Tajima's $D = -1.02$, Fu and Li's $D = 0.12$, Fu and Li's $F = -0.27$). These results suggest that demographic factors in the beetle population are probably not responsible for the patterns observed in the mtDNA data set, as demography is expected to affect all the genes in the genome. Factors that may confound this comparison are considered in the discussion.

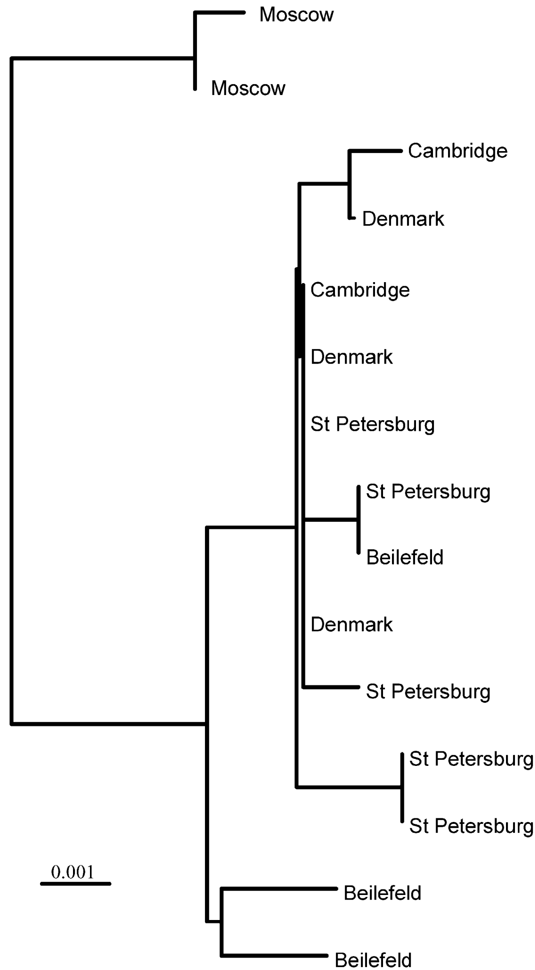

FIGURE 2.—Neighbor-joining tree of *g6pd* sequences rooted with a sequence from *Adalia decempunctata* (not shown).

We tested whether there was an excess of polymorphism in the mtDNA data set, using an HKA test (HUDSON *et al.* 1987). Under the neutral hypothesis, the level of intraspecific polymorphism and interspecific divergence will be positively correlated across different loci in the genome. However, balancing selection on cytoplasmic genes will result in elevated intraspecific mtDNA polymorphism. We tested this hypothesis by comparing intraspecific polymorphism within *A. bipunctata* and interspecific divergence from *A. decempunctata* in the nuclear gene *g6pd* and mtDNA. The level of polymorphism in the mtDNA was significantly higher than expected under neutrality (Table 3; $\chi^2 = 9.2$, $P < 0.005$).

This test is highly conservative, as we assumed that the effective population size of *g6pd* is three times that of the mtDNA linked to the Rickettsia bacterium. This would be the case if all females were infected, because cytoplasmic genes are haploid and uniparentally inherited. However, the effective population size of mtDNA in *A. bipunctata* is approximately the number of infected females, not the number of females (JOHNSTONE and HURST 1996). This is because imperfect transmission of the bacterium results in the rapid replacement of mitochondrial haplotypes in uninfected females with those from infected females. The average prevalence is <5%, giving an effective population size in the order of 1–2% that of *g6pd*. Therefore, the true expected level of polymorphism is substantially less than that shown in Table 3.

TABLE 3

HKA test comparing mitochondrial DNA and *g6pd*

	mtDNA	*g6pd*
No. of sites	1173	1216
Intraspecific polymorphism		
No. of alleles	13	15
Segregating sites	78 (46.95)	16 (47.05)
Interspecific divergence		
Mean no. of differences	123.85 (154.90)	78.93 (47.89)

Expected values are given in parentheses.

Population subdivision will increase the level of within-species diversity and can therefore confound an HKA test. This problem is greatest if the loci being compared have very different levels of population subdivision, which could occur in our case if males and females have different migration rates (INGVARSSON 2004). In the DISCUSSION we consider this from a natural history perspective and conclude that males and females have similar patterns of dispersal. Here, we correct our data for this effect. If it is assumed that the population structure of this species conforms to WRIGHT's (1930) island model, the data can be corrected for the difference in structure between the two loci (INGVARSSON 2004). We estimated F_{ST} (WEIR and COCKERHAM 1984) from our mitochondrial data (Table 1) to be 0.315. This was little changed if the Scandinavian populations are pooled. We conservatively assumed that the nuclear locus lacked any population structure. We then increased the intraspecific polymorphism and decreased interspecific divergence by the factor $(1 - F_{ST})$ (INGVARSSON 2004). The HKA test remained significant ($\chi^2 = 5.03$, $P = 0.025$).

The significant HKA test result could potentially be produced by a selective sweep at the *g6pd* locus rather than by elevated cytoplasmic variation. However, we were unable to detect any departures from neutrality in the *g6pd* data set (see above), suggesting that this is not the case. In addition to the analyses described above, we used the McDonald-Kreitman (MCDONALD and KREITMAN 1991) test to compare the numbers of synonymous and nonsynonymous mutations that were polymorphic within the population and fixed between species, and no significant deviation was found (synonymous fixed = 69, synonymous polymorphic = 15, nonsynonymous fixed = 7, nonsynonymous polymorphic = 1: not significant).

Low diversity within a mtDNA clade: The results above indicate that natural selection is maintaining

three different mitochondrial haplotypes within this species. Next, we examined the evolutionary processes occurring within one of these clades. We sequenced a longer stretch of mtDNA from nine infected females from the A clade. We obtained 3941 bp of sequence from each of six individuals from Cambridge in the UK, from two individuals from Malmo in Scandinavia, and from one individual from Ljusdal in Scandinavia. The diversity of these sequences was very low (π = 0.00006), with only a single derived mutation being detected in one of the Cambridge females.

These data can be used to estimate the time since the haplotypes shared a common ancestor. If we assume both that the low diversity results from a recent selective sweep or population bottleneck and that the population size is large, then the genealogy of the sequences will be star shaped. Therefore, the time t since they coalesce to a common ancestor will be a function of the number of polymorphic sites S, the per site per year mutation rate μ, the number of individuals sequenced n, and the length of those sequences l:

$$t = \frac{S}{\mu n l}.$$

Assuming that the average value of μ across all sites is 1.15×10^{-8} (Brower 1994), then these cytotypes shared a common ancestor ~2500 years ago (95% confidence interval assuming S is poisson distributed: 127–14,100 years). If the genealogy of these sequences is not star shaped, then the true confidence limits may be wider. *A. bipunctata* in Cambridge usually has only one generation per year, but in some years a proportion go through a second generation (Majerus 1994). Therefore, 2500 years equates to ~3000 generations.

Mitochondrial variation in uninfected females: Finally, we compared the frequencies of the three haplotypes in the infected and uninfected females. Under the null hypothesis that the three bacterial and mitochondrial haplotypes are biologically equivalent and at equilibrium, the frequencies of mtDNA haplotypes in the infected and uninfected females will be the same. This is because infected females produce some uninfected daughters, resulting in unidirectional mtDNA gene flow from infected to uninfected females (Johnstone and Hurst 1996; Jiggins 2003). We were able to reject this null hypothesis, as there are large differences in haplotype frequency between infected and uninfected females within the same population (Table 1). In 1995 the A haplotype occurred in 100% of infected Cambridge females, but in only 67% of uninfected females (Monte Carlo simulation, $P < 0.00001$). Similarly, in 2003 the A haplotype occurred in 96% of infected females, but in only 53% of uninfected females (Monte Carlo simulation, $P < 0.00005$).

If, as predicted by theory, the mtDNA in uninfected females is recently derived from infected females then we expected the mtDNA sequences in uninfected females to be very similar to those in infected females. The 800 bp of mtDNA we sequenced from 10 A-haplotype females was identical to the mtDNA sequenced below from A-haplotype infected beetles. Therefore, although the frequency of the mtDNA haplotypes differs between infected and uninfected females, their sequences do not.

DISCUSSION

High cytoplasmic diversity: Rickettsia-infected *A. bipunctata* beetles have very high levels of mtDNA variation. This is in striking contrast to previous studies of male killers and other similar bacteria, which are typically associated with extremely low levels of mitochondrial variation (Ballard *et al.* 1996; Jiggins 2003). In these species, it is thought that a single mtDNA haplotype has hitchhiked through the population following the recent invasion of the bacterium. Therefore, these previous studies suggest that the infections were short lived within host populations and frequently switched host species. Our results demonstrate that this is not always the case, and polymorphisms of male-killing bacteria can persist for long periods of time. Mitochondrial DNA is widely used as a marker in demographic studies, and it is well known that selection on cytoplasmic symbionts can confound such analyses (Hurst and Jiggins 2005). Our results show that these effects can be considerably more complex than simple reductions in diversity following a selective sweep.

Not only was the mtDNA from infected beetles highly variable, but also in our sample this variation was partitioned into only three haplotypes. This unusual pattern could be generated by either demographic processes or natural selection. We attempted to separate these effects by also sequencing a nuclear gene, as demographic processes will affect the whole genome, while selection will affect only cytoplasmic genes. The nuclear gene showed no departures from the neutral model and was significantly less diverse than the mtDNA, suggesting that selection may be maintaining variation in the cytoplasmic genome.

This comparison between nuclear and cytoplasmic genes could potentially be confounded if females have much lower migration rates than males (Ingvarsson 2004). However, from our knowledge of the natural history of ladybirds, any sex differences are likely to be small. Long-distance dispersal of both sexes occurs to and from over-wintering sites or following crashes in prey numbers (Majerus 1994). More local movements occur between host plants, and here females migrate somewhat more than males, perhaps due to their longer lifespan (Brakefield 1984; Tinsley 2003). In a more formal approach to this problem, the HKA test was still significant after correction for population subdivision, despite several conservative assumptions.

The population genetic tests that we have performed do not distinguish selection acting directly on the mtDNA from selection acting on the bacterium. This is a problem common to any population genetic analysis in a nonrecombining region. There are, however, reasons to believe that balancing selection is more likely to be acting on the male killer than on mitochondrial genes. Selection on the mitochondrion would affect mtDNA variation in infected and uninfected females equally. However, there are dramatic differences between the mitotype frequencies in infected and uninfected females, suggesting that natural selection is not solely acting on the mtDNA itself. Furthermore, aside from unusual situations such as self-incompatibility in plants, evidence for selection maintaining long-term polymorphisms is comparatively scarce outside host-parasite systems such as this. It may be informative to gain more sequence data from the bacteria themselves. For example, we have no data on whether or not the bacteria associated with the A and C haplotypes are different.

Why might selection maintain multiple male-killer strains? The first hypothesis is that the fitness of male killers is frequency dependent. Theory predicts that frequency dependence can result if there are costly host genes that confer resistance to specific male killers (Randerson *et al.* 2000). Models of host genes that confer resistance against specific parasite (Barrett 1988) or sex-ratio-distorter (Frank 1989) genotypes also predict the geographical variation that we observed. This hypothesis predicts that host populations will contain polymorphic genes that are costly and confer resistance against specific Rickettsia genotypes. Although resistance genes against male killers have been described in Drosophila, there are no data on whether they target specific bacterial genotypes or whether similar genes exist in Adalia (Cavalcanti *et al.* 1957; Malogolowkin 1958). A second hypothesis is that the polymorphism is maintained by a selection-migration balance. Under this scenario different Rickettsia strains are adapted to different populations (*e.g.*, to different temperature regimes, host life cycles, or host physiologies). If this is coupled to migration between those populations, then a long-term polymorphism could result. Our data do not distinguish between these two hypotheses.

Geographical and temporal variation: A striking aspect of our data is the large geographical variation in the frequency of the three cytotypes. There is also temporal variation in cytotype frequency in uninfected females, although our samples were too small to be certain whether similar changes occur in infected females. Therefore, the polymorphism does not seem to have settled to a stable equilibrium frequency across populations. This contrasts with a previous study of a male killer where a recent selective sweep erased population structure from mitochondrial but not nuclear genes (Jiggins 2003).

The highly spatial structure of mtDNA could be the result of either selection on cytoplasmic genes (frequency dependence or local adaptation) or simply drift and demography in a subdivided population. Microsatellites in this species show far lower levels of population subdivision than we have observed for mtDNA (Haddrill 2001). However, the small effective population size of mtDNA in the presence of a male killer (Johnstone and Hurst 1996) may explain this discrepancy. Therefore, it will be necessary to test whether the relative fitness of strains differs between populations to separate these hypotheses.

The highly structured male-killer population may play a role in stabilizing the polymorphism. If a particular male-killer strain is lost from a local population it is unlikely to be lost from the entire metapopulation, so it can subsequently reappear due to immigration (Judson 1995). The importance of this process is supported by the observation that in some populations certain strains either have been lost or are extremely rare.

Similar patterns of high mitochondrial diversity and geographical variation have been found in studies of cytoplasmic male sterility (CMS) in plants. CMS causes normally hermaphroditic plants to develop into females because they carry mitochondria that prevent the production of pollen. It is analogous to male killing as it releases more resources for seed production and may reduce inbreeding depression by preventing self-fertilization (van Damme 1984; Sun and Ganders 1986; Kohn 1988). Plant species are often host to multiple CMS haplotypes that can vary greatly in their relative frequencies between populations (reviewed in Frank 1989). In the gynodioecious plant *Silene acaulis,* this diversity of mitochondrial haplotypes has persisted for >15 million years (Stadler and Delph 2002). These patterns have been interpreted as being the product of coevolution between the mitochondria causing male sterility and nuclear genes restoring male fertility. As we lack any data on the existence and nature of nuclear resistance genes in this species, it is unclear whether similar processes underlie the patterns we have seen.

Low within-genotype diversity: The three Rickettsia-associated cytotypes contain very little genetic variation despite having persisted for >2 million years. For example, isolates of the A strain from across Europe coalesce to a common ancestor just 2500 years ago. The low diversity of this genotype may be the result of selection. For example, cycles in strain prevalence will result in the loss of diversity as each genotype passes though periodic bottlenecks (Stahl *et al.* 1999).

The low diversity of this bacterial clade could also reflect its small population size. We are unable to reject this hypothesis, but consider it to be less likely than explanations relying on selection. It would require the effective population size of infected females to be just a few thousand on the basis of the estimates of mtDNA diversity and mutation rate described above. Given that *A. bipunctata* is an extremely common insect with normal levels of nuclear diversity, and 11% of Cambridge

females are infected with this strain (Table 1), this figure is extremely small.

Uninfected females: The mitochondrial haplotype frequencies in infected and uninfected females are very different. This is noteworthy because under the null model that the three strains and their associated mitochondria have identical biological properties, equilibrium mtDNA frequencies in infected and uninfected females will be the same (Jiggins 2003). There are several possible explanations of why this is not the case, but all rely on fitness differences between the three bacterial strains. First, the system may not be at equilibrium because the relative frequency of the Rickettsia haplotypes has recently changed due to natural selection. Second, selection may favor different Rickettsia strains in different populations, while uninfected beetles migrate between those populations. Third, the strains may have different rates of transmission from mother to daughter (and transmission efficiency is an important component of fitness). Finally, the fittest male killer may be associated with a less fit mitochondrion. Therefore, this mtDNA type will be underrepresented in the uninfected population relative to the infected population.

In conclusion, male-killing Rickettsia are associated with high levels of mitochondrial polymorphism. The data suggest that this variation has been maintained by natural selection. This generates the testable prediction that the relative fitness of the Rickettsia strains will vary with environmental conditions or host genetic background.

Greg Hurst, Hinrich Schulenburg, and Mike Majerus all assisted with collecting beetles. We particularly thank Anna Corry for the 1995 sample. We thank two anonymous reviewers and the editor for incisive comments that greatly improved this manuscript. The authors are supported by the Wellcome Trust and Natural Environment Research Council.

LITERATURE CITED

Ballard, J. W. O., J. Hatzidakis, T. L. Karr and M. Krietman, 1996 Reduced variation in *Drosophila simulans* mitochondrial DNA. Genetics **144:** 1519–1528.

Barrett, J. A., 1988 Frequency-dependent selection in plant fungal interactions. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. **319:** 473–483.

Brakefield, P. M., 1984 Ecological-studies on the polymorphic ladybird Adalia- bipunctata in The Netherlands. 1. Population biology and geographical variation of melanism. J. Anim. Ecol. **53:** 761–774.

Brower, A. V. Z., 1994 Rapid morphological radiation and convergence among races of the butterfly Heliconius-erato inferred from patterns of mitochondrial-DNA evolution. Proc. Natl. Acad. Sci. USA **91:** 6491–6495.

Cavalcanti, A. L. G., D. N. Falcao and L. E. Castro, 1957 "Sex-ratio" in *Drosophila prosaltans*—a character due to interaction between nuclear genes and cytoplasmic factors. Am. Nat. **91:** 327–329.

Charlesworth, D., 1981 A further study of the problem of the maintenance of females in gynodioecious species. Heredity **46:** 27–39.

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

Davis, M. J., Z. T. Ying, B. R. Brunner, A. Pantoja and F. H. Ferwerda, 1998 Rickettsial relative associated with papaya bunchy top disease. Curr. Microbiol. **36:** 80–84.

Depaulis, F., and M. Veuille, 1998 Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol. Biol. Evol. **15:** 1788–1790.

Depaulis, F., S. Mousset and M. Veuille, 2001 Haplotype tests using coalescent simulations conditional on the number of segregating sites. Mol. Biol. Evol. **18:** 1136–1138.

Frank, S. A., 1989 The evolutionary dynamics of cytoplasmic male sterility. Am. Nat. **133:** 345–376.

Frank, S. A., 1998 Dynamics of cytoplasmic incompatibility with multiple Wolbachia infections. J. Theor. Biol. **192:** 213–218.

Fu, Y.-X., and W.-H. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

Gouyon, P. H., F. Vichot and J. M. M. Vandamme, 1991 Nuclear-cytoplasmic male-sterility—single-point equilibria versus limit-cycles. Am. Nat. **137:** 498–514.

Haddrill, P. R., 2001 The development and use of molecular genetic markers to study sexual selection and population genetics in the 2-spot ladybird, Adalia bipunctata (L.). Ph.D. Thesis, Department of Genetics, University of Cambridge, Cambridge, UK.

Hudson, R. R., M. Kreitman and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Hurst, G. D. D., and F. M. Jiggins, 2005 Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc. R. Soc. Lond. Ser. B **272:** 1525–1534.

Hurst, G. D. D., and M. E. N. Majerus, 1993 Why do maternally inherited microorganisms kill males? Heredity **71:** 81–95.

Hurst, G. D. D., M. E. N. Majerus and L. E. Walker, 1992 Cytoplasmic male killing elements in *Adalia bipunctata* (Linnaeus) (Coleoptera: Coccinellidae). Heredity **69:** 84–91.

Hurst, G. D. D., F. M. Jiggins, J. H. G. Schulenberg, D. Bertrand, S. A. West *et al.*, 1999a Male-killing *Wolbachia* in two species of insect. Proc. R. Soc. Lond. Ser. B **266:** 735–740.

Hurst, G. D. D., J. H. G. v. d. Schulenburg, T. M. O. Majerus, D. Bertrand, I. A. Zakharov *et al.*, 1999b Invasion of one insect species, *Adalia bipunctata*, by two different male-killing bacteria. Insect Mol. Biol. **8:** 133–139.

Hurst, L. D., 1991 The incidences and evolution of cytoplasmic male killers. Proc. R. Soc. Lond. Ser. B **244:** 91–99.

Hurst, L. D., 1993 The incidences, mechanisms and evolution of cytoplasmic sex ratio distorters in animals. Biol. Rev. **68:** 121–193.

Iablokoff-Khnzorian, S. M., 1982 *Les Coccinelles. Coléoptères-Coccinellidae. Tribu Coccinellini des régions Palearctique et Orientale.* Société Nouvelle des Éditions Boubée, Paris.

Ingvarsson, P. K., 2004 Population subdivision and the Hudson-Kreitman-Aguade test: testing for deviations from the neutral model in organelle genomes. Genet. Res. **83:** 31–39.

Jiggins, F. M., 2003 Male-killing Wolbachia and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. Genetics **164:** 5–12.

Johnstone, R. A., and G. D. D. Hurst, 1996 Maternally inherited male-killing microorganisms may confound interpretation of mtDNA variation in insects. Biol. J. Linn. Soc. **53:** 453–470.

Judson, O. P., 1995 Preserving genes—a model of the maintenance of genetic- variation in a metapopulation under frequency-dependent selection. Genet. Res. **65:** 175–191.

Jukes, T. H., and C. R. Cantor, 1969 Evolution of protein molecules, pp. 21–132 in *Mammalian Protein Metabolism*, edited by H. N. Munro. Academic Press, New York.

Kohn, J. R., 1988 Why be female. Nature **335:** 431–433.

Kreitman, M., and R. R. Hudson, 1991 Inferring the evolutionary histories of the Adh and Adh-Dup loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics **127:** 565–582.

Laser, K. D., and N. R. Lersten, 1972 Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. Bot. Rev. **38:** 425–454.

Lewontin, R. C., and J. Felsenstein, 1965 The robustness of homogeneity tests in 2 X n tables. Biometrics **21:** 19–33.

Lus, Y. Y., 1947 Some aspects of population increase in *Adalia bipunctata* l. 2. the strains without males. Dokl. Nauk SSR **57:** 951–954.

Majerus, M. E. N., 1994 *Ladybirds*. Harper Collins, London.

Malogolowkin, C., 1958 Maternally inherited "sex ratio" condition in *Drosophila willistoni* and *D. paulistorum.* Genetics **43:** 274–286.

Markovtsova, L., P. Marjoram and S. Tavaré, 2001 On a test of Depaulis and Veuille. Mol. Biol. Evol. **18:** 1132–1133.

McDonald, J. H., and M. Kreitman, 1991 Adaptive Protein Evolution at the Adh Locus in Drosophila. Nature **351:** 652–654.

Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Randerson, J. P., N. G. C. Smith and L. D. Hurst, 2000 The evolutionary dynamics of male-killers and their hosts. Heredity **84:** 152–160.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu *et al.*, 1994 Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. **87:** 651–701.

Stadler, T., and L. F. Delph, 2002 Ancient mitochondrial haplotypes and evidence for intragenic recombination in a gynodioecious plant. Proc. Natl. Acad. Sci. USA **99:** 11730–11735.

Stahl, E. A., G. Dwyer, R. Mauricio, M. Kreitman and J. Bergelson, 1999 Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. Nature **400:** 667–671.

Sun, M., and F. R. Ganders, 1986 Female frequencies in gynodioecious populations correlated with selfing rates in hermaphrodites. Am. J. Bot. **73:** 1645–1648.

Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.

Takahata, N., and M. Nei, 1990 Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics **124:** 967–978.

Tinsley, M., 2003 The ecology and evolution of male-killing bacteria in ladybirds. Ph.D. Thesis, Department of Genetics, University of Cambridge, Cambridge, UK.

van Damme, J. M. M., 1984 Gynodioecy in *Plantago lanceolata* L. III. Sexual reproduction and the maintenance of male steriles. Heredity **52:** 77–93.

Vekemans, X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137:** 1157–1165.

von der Schulenburg, J. H. G., M. Habig, J. J. Sloggett, K. M. Webberley, D. Bertrand *et al.*, 2001 Incidence of male-killing *Rickettsia* spp. (a-proteobacteria) in the ten-spot ladybird beetle *Adalia decempunctata* L. (Coleoptera: Coccinellidae). Appl. Environ. Microbiol. **67:** 270–277.

von der Schulenburg, J. H. G., G. D. D. Hurst, D. Tetzlaff, G. E. Booth, I. A. Zakharov *et al.*, 2002 History of infection with different male-killing bacteria in the two-spot ladybird beetle *Adalia bipunctata* revealed through mitochondrial DNA analysis. Genetics **160:** 1075–1086.

Wall, J. D., and R. R. Hudson, 2001 Coalescent simulations and statistical tests of neutrality. Mol. Biol. Evol. **18:** 1134–1135.

Walsh, P. S., D. A. Metzger and R. Higuchi, 1991 Chelex 100 as a medium for simple extraction of DNA for PCR based typing of forensic material. Biotechniques **10:** 506–513.

Watterson, G. A., 1975 On the number of segregating sites in models without recombination. Theor. Popul. Biol. **7:** 256–276.

Weir, B. S., and C. C. Cockerham, 1984 Estimating F-statistics for the analysis of population-structure. Evolution **38:** 1358–1370.

Werren, J. H., 1987 The coevolution of autosomal and cytoplasmic sex ratio factors. J. Theor. Biol. **124:** 317–334.

Werren, J. H., G. D. D. Hurst, W. Zhang, J. A. J. Breeuwer, R. Stouthamer *et al.*, 1994 Rickettsial relative associated with male killing in the ladybird beetle (*Adalia bipunctata*). J. Bacteriol. **176:** 388–394.

Williams, S. G., J. B. Sacci, M. E. Schriefer, E. M. Andersen, K. K. Fujioka *et al.*, 1992 Typhus and typhus-like Rickettsiae associated with opossums and their fleas in Los-Angeles County, California. J. Clin. Microbiol. **30:** 1758–1762.

Wright, S., 1930 Evolution in Mendelian populations. Genetics **16:** 97–159.

Communicating editor: M. Uyenoyama