Increased Rates of Sequence Evolution in Endosymbiotic Bacteria and Fungi with Small Effective Population Sizes

Megan Woolfit and Lindell Bromham

Centre for the Study of Evolution, School of Biological Sciences, University of Sussex, United Kingdom

Mutualistic, maternally transmitted endosymbiotic microorganisms undergo severe population bottlenecks at each host generation, resulting in a reduction in effective population size (N_e). Previous studies of *Buchnera*, the primary endosymbiont of aphids, and of several other species of endosymbiotic bacteria have shown that these species exhibit an increase in the rate of substitution of slightly deleterious mutations, among other predicted effects of increased drift due to small N_e, such as reduced codon bias. However, these studies have been limited in taxonomic scope, and it was therefore not clear whether the increase in rate is a general feature of endosymbiont lineages. Here, we test the prediction that a long-term reduction in N_e causes an increase in substitution rate using DNA sequences of the 16S rRNA gene from 13 phylogenetically independent comparisons between taxonomically diverse endosymbiotic microorganisms and their free-living relatives. Maximum likelihood and distance-based methods both indicate a significant increase in substitution rate in a wide range of bacterial and fungal endosymbionts compared to closely related free-living lineages. We use the same data set to test whether 16S genes from endosymbionts display increased A + T content, another indicator of increased genetic drift, and find that there is no significant difference in base composition between endosymbiont and nonendosymbiont 16S genes. However, analysis of an additional data set of whole bacterial genomes demonstrates that, while host-dependent bacteria have significantly increased genomic A + T content, the base content of the 16S gene tends to vary less than that of the whole genome. It is possible that selection for stability of rRNA is strong enough to overcome the effects of drift toward increased A + T content in endosymbiont 16S genes, despite the reduced effective population sizes of these organisms.

Introduction

Effective population size (N_e) is a reflection of the strength of stochastic processes in the evolution of a population. It affects many evolutionary parameters, including the rate at which neutral genetic variation is lost, the efficiency of selection, and the rate of accumulation of nearly neutral mutations (Kimura and Ohta 1971; Ohta 1992). Nearly neutral mutations are those whose selective coefficients (s) are small enough that they may behave in a small population (where $|N_e s| < 1$) as if they were strictly neutral (Ohta 1972). As effective population size decreases, the number of mutations whose selective coefficients are less than or equal to $1/N_e$ will increase, and thus a higher proportion of mutations will be nearly neutral. Slightly deleterious mutations will then have an increased likelihood of drifting to fixation in the population rather than being eliminated by selection, thus increasing the rate of substitution (Ohta 1972, 1992). Ohta's original formulation of this theory was based on a model of irreversible fixation of deleterious mutations, but these predictions also hold for reversible models under equilibrium conditions (Gillespie 1994). Under either class of model, if nearly neutral mutations make up a significant proportion of total mutations in a population, as has been suggested by several recent studies (reviewed in Eyre-Walker et al. 2002), then the increased substitution rate in small populations should be detectable in analyses of sequence data.

Two early tests of this prediction, using sequence data from organisms assumed to differ in population size, found that species with smaller N_e had faster rates of sequence evolution (Wu and Li 1985; DeSalle and Templeton 1988). However, each analysis was based on a single comparison, which limits the conclusions which can be drawn from them, a problem shared by several more recent studies (Easteal and Collet 1994; Kliman et al. 2000; Weinreich 2001). Spradling, Hafner, and Demastes (2001) compared data from 21 species of rodents and found an inverse relationship between N_e and rate of evolution, but their analysis did not use phylogenetically independent comparisons (see below). Johnson and Seger's (2001) study compared rates of evolution in mainland and island species of birds and found an increase in substitution rate in island species, which are assumed to have restricted population sizes. However, many of their comparisons overlapped on the phylogeny, and so they provide relatively few statistically independent data points to test the hypothesis.

Endosymbiotic microorganisms provide a further test of the effect of N_e on substitution rates (Moran 1996). These mutualistic bacteria and fungi live sequestered within the cells of eukaryotic hosts, and are maternally transmitted by the infection of eggs or embryos (Buchner 1965). This mode of transmission results in endosymbiotic lineages remaining sequestered within a host lineage, with strictly vertical transmission, dividing the population into isolated subpopulations which do not exchange genetic material. Relatively few endosymbionts are thought to be transmitted in each inoculation—for various *Buchnera* species, the number is estimated to be between 40 and several thousand individuals (Buchner 1965; Rispe and Moran 2000; Mira and Moran 2002), a small percentage of the population found in each adult host (Mira and Moran 2002). Although bottleneck size has not been quantified for other endosymbiont species, it seems likely that many of them undergo a similar degree of population reduction. These two characteristics of endosymbiont biology, population subdivision and recurrent severe reductions in population size

Key words: symbiosis, yeast-like symbionts, phylogeny, AT bias, relative rate, likelihood ratio test.

E-mail: m.r.q.woolfit@sussex.ac.uk.

Mol. Biol. Evol. 20(9):1545–1555. 2003
DOI: 10.1093/molbev/msg167
Molecular Biology and Evolution, Vol. 20, No. 9,

(bottlenecks), both act to reduce the effective population size of these species (Rispe and Moran 2000).

Asexuality is also expected to reduce effective population size, because of the increased linkage of the genome (Felsenstein 1974; Pamilo, Nei, and Li 1987; Lynch and Blanchard 1998). It is not known whether bacterial endosymbionts are sexual, although it is probable that *Buchnera* at least is largely clonal, as it has lost several genes that facilitate recombination, including *recA* (Tamas et al. 2002). Vertically transmitted endosymbionts are in any case effectively clonal, as any recombination can only occur between very closely related individuals.

These factors—population subdivision, recurrent bottlenecks and effective asexuality—are expected to lead to an increased rate of fixation of slightly deleterious mutations within endosymbiont lineages. Previous analyses of sequence data from the endosymbionts of aphids (*Buchnera*) and psyllids (*Candidatus* Carsonella) have shown that, as predicted, these bacteria exhibit higher rates of substitution than free-living bacteria (Moran 1996; Thao et al. 2000), as well as other predicted effects of increased drift, such as lack of adaptive codon bias and greatly increased genomic A + T content (Clark et al. 1992; Brynnel et al. 1998; Wernegreen and Moran 1999). Genomic base composition reflects a balance between mutational biases and selection for translational efficiency. Species with large effective population sizes are expected to exhibit strong bias toward optimal codon usage. In smaller populations, however, in which translational selection is weaker, patterns of codon usage largely reflect drift and mutational biases, as has been shown in *Buchnera* and *Candidatus* Carsonella (e.g. Spaulding and von Dohlen 1998; Wernegreen and Moran 1999).

The cause of a mutational bias toward A + T in *Buchnera* and *Candidatus* Carsonella is not known, although it has been suggested that host dependence itself might result in an A + T bias (Rocha and Danchin 2002). In an analysis based on 50 bacterial whole genomes (which did not take into account phylogenetic nonindependence) it was found that average genomic base composition differed significantly between host-dependent (62% A + T) and free-living or facultatively pathogenic bacteria (51% A + T). Their explanation for this, that host-dependent bacteria must compete with their host for scarce metabolic resources, and thus experience a reduced availability of metabolically more expensive cytosine and guanine, offers a possible mechanism for the A + T mutation bias observed in these two endosymbiont species.

Reports of faster substitution rates, and in some cases other indicators of increased genetic drift, in endosymbionts of aphids, psyllids, whiteflies, tsetse flies, and mealybugs (Moran 1996; Spaulding and von Dohlen 2001) suggest that reduced N_e may affect the rate and pattern of molecular evolution in these species. However, all but one of the mutualistic endosymbionts previously examined are γ-Proteobacteria, and several are very closely related (Clark et al. 1992; Munson, Baumann, and Moran 1992; Spaulding and von Dohlen 1998). Therefore it is not clear whether the increased substitution rate detected is due to reduced effective population size caused by the endosymbiotic lifestyle, or to some other influence on rate in this lineage of γ-Proteobacteria. To tease apart possible explanations, it is necessary to analyze data from many more groups of endosymbionts, and to take into account information on phylogenetic relationships.

Species cannot be treated as independently derived data points because they form part of a phylogenetic hierarchy. Therefore similarity between species can vary according to their level of relatedness, rather than randomly, as would be expected if species traits were independently acquired (Harvey and Pagel 1991). Cross-species correlations of traits which do not account for this nonindependence are subject to inflated type I and type II errors and may be misleading (Martins and Garland 1991; Harvey and Rambaut 1998). The method of phylogenetically independent contrasts (Felsenstein 1985) corrects for species nonindependence by using a phylogeny to identify a set of independent comparisons between pairs of species, nodes, or clades which differ in the trait of interest (e.g., endosymbiosis). Any difference in other traits between the two species, nodes, or clades (e.g., rate of molecular evolution) must have arisen since their last common ancestor, and independently of any other pair. Each comparison then contributes one independent data point to a statistical analysis. The importance of using phylogenetically independent comparisons was recognized in Moran's (1996) original analysis of rate acceleration in endosymbionts; however, data were available for only a few comparisons.

It is important to note that an assumption of independent origin of endosymbiosis is not sufficient to consider comparisons *phylogenetically* independent. Endosymbiosis may have arisen four or more times in the Enterobacteriaceae, but because there is only one known sister group for all four lineages (see fig. 1, γ-Proteobacteria 5), we cannot exclude the possibility that an event which affected the rate of molecular evolution (mutation of a gene coding for an error-correcting enzyme, for example) could have occurred on the branch ancestral to this group, causing a change in evolutionary rate in all endosymbiotic lineages within it. In this case it would be misleading to count each of these four lineages as an independent data point in terms of rate change, even though they probably do represent four independent transitions to endosymbiosis (Buchner 1965; Lambert and Moran 1998).

The analysis presented here employs the large volume of 16S ribosomal RNA (16S) sequence data from bacterial and fungal endosymbionts which has become available in public databases in the last few years. Sequence data from six major groups of bacteria and one class of fungi are included, giving a much wider taxonomic range than previous studies. To test whether endosymbionts have higher rates of substitution, we apply two different methods of rate estimation to 13 phylogenetically independent comparisons, each consisting of a clade of endosymbionts and their closest nonendosymbiotic relatives. We also use this data set to test whether endosymbiont 16S genes exhibit higher A + T content, another indicator of genetic drift from reduced N_e. Finally, we apply the comparative method to an analysis of bacterial whole genomes to determine whether the relationship between host dependence and A + T content

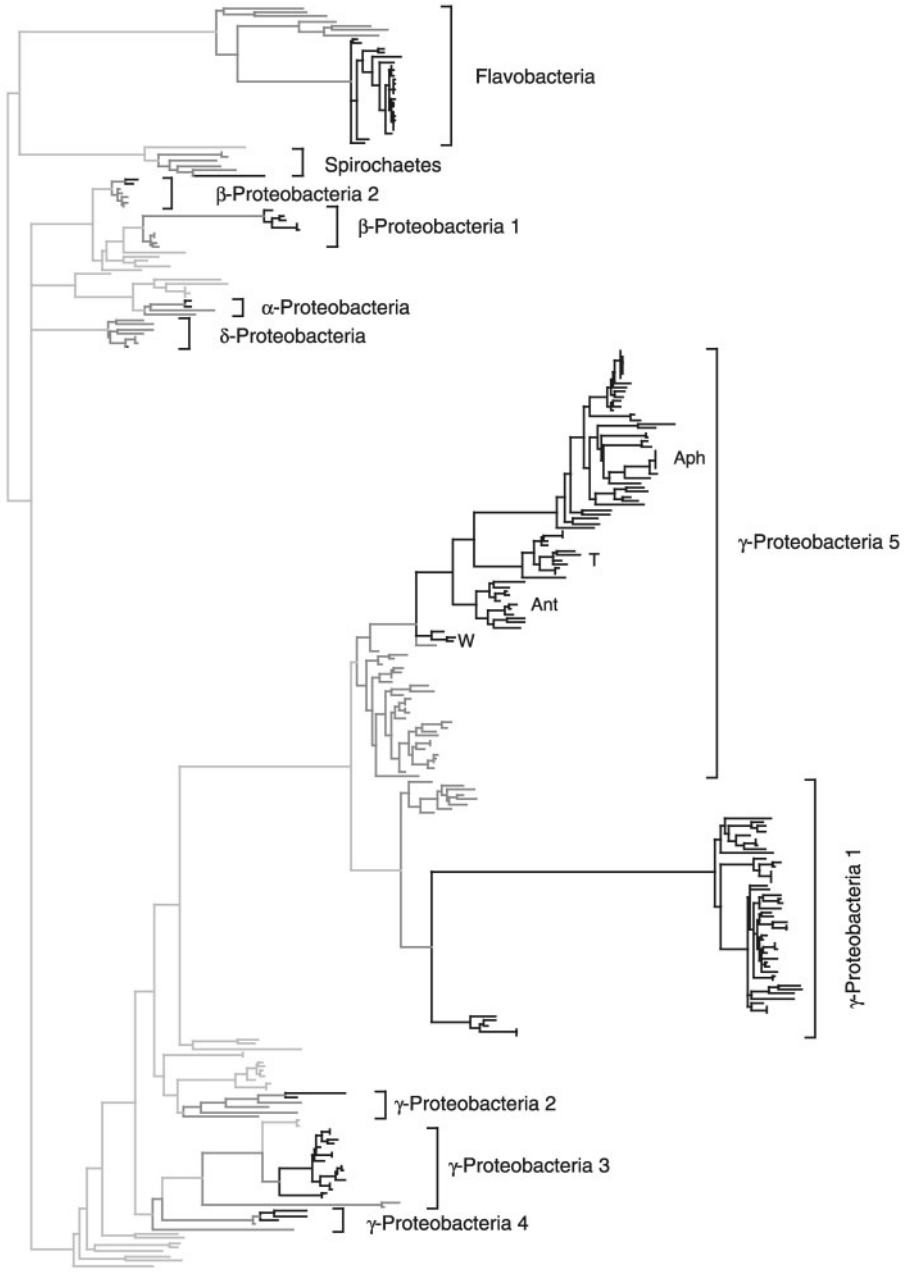

FIG. 1.—Maximum likelihood (ML) trees based on 16S sequences, including all species from the bacterial independent comparisons. Labeled brackets indicate comparisons, each consisting of a clade of endosymbiotic species and their closest nonendosymbiotic relatives (see table 1). The spine of the tree is light gray, branches leading to nonendosymbiont species in each comparison are in dark gray, and those leading to endosymbiont species are black. Numbers (e.g., γ-Proteobacteria 1) refer to separate comparisons made within a taxonomic group. The Enterobacteriaceae (γ-Proteobacteria 5) comprises four lineages of endosymbionts, labeled by host group: Aph = aphids, T = tsetse flies, Ant = carpenter ants, W = weevils.

noted by Rocha and Danchin (2002) is still observed when phylogenetic nonindependence is taken into account. We then test whether the base composition of the 16S genes of these species accurately reflects that of the whole genome.

Methods

Data

Sequence data for the 16S gene were obtained from GenBank (available online at http://www.ncbi.nlm.nih.gov/Genbank/index.html) for all of the vertically transmitted endosymbionts represented in the database. Multiple isolates of species were included if they were present, resulting in a total of 202 sequences. To maximize the number of independent comparisons, BLAST searches were used to identify close nonendosymbiotic relatives of each endosymbiont group. Where possible, completely free-living species were chosen to maximize the difference in effective population size between endosymbionts and nonendosymbionts, although facultative pathogens were used when necessary.

Phylogenetic Analysis

Sequences were selected from six major groups of bacteria (the α-, β-, δ-, and γ-Proteobacteria, Flavobacteria, and Spirochaetes) and one group of fungi (Ascomycetes). Each of these seven data sets was aligned and analyzed separately. Sequences were aligned by eye using Se-Al (Rambaut 2002). Each alignment consisted of approximately 1,600 bp. Regions of sequence which could not be aligned unambiguously (typically <120 bp in total) were excluded from the analysis. For each of the seven data sets, phylogenies were constructed in PAUP* (Swofford 2002) using two methods, maximum likelihood (ML) and LogDet. Maximum likelihood phylogenies were constructed with an HKY + Γ model of nucleotide substitution (Hasegawa, Kishino, and Yano 1985; Yang 1994) with the transition-transversion ratio (ti/tv) and the gamma shape parameter (α) estimated from the data. The HKY model of substitution allows base frequencies of the four nucleotides to vary, but it assumes that these frequencies are approximately constant across the sequences being compared. Because some endosymbiont sequences have very biased base compositions (Clark, Moran, and Baumann 1999), and this can affect topology reconstruction (Mooers and Holmes 2000), a second model, LogDet, was also used. The LogDet transformation is robust to variations in base composition across the tree (Lockhart et al. 1994).

The topologies of the maximum likelihood and LogDet trees were compared and phylogenetically independent comparisons, each consisting of a monophyletic group of endosymbionts and a closely related group of nonendosymbionts, were chosen only if they were supported by both trees (see fig. 1 and table 1 for these comparisons). Although some comparisons differ in the number of species in the endosymbiotic and nonendosymbiotic groups, this is not expected to bias our estimation of rates. Maximum likelihood methods, unlike parsimony (Sanderson 1990), explicitly model multiple substitutions per site and are therefore expected to be robust to ''node density effect'' (Bromham et al. 2002). Both methods of rate estimation require outgroups; we chose the most closely related definite outgroup for each comparison.

Estimating Substitution Rates

For each independent comparison, rates of nucleotide substitution in endosymbiont and nonendosymbiont groups were estimated and compared using two methods: a likelihood ratio test and a distance-based relative rate test (table 2). For each method, a comparison was scored as positive if the endosymbiont lineage had a higher rate of substitution than the free-living lineage, and each comparison contributed one point to a signed-ranks test. This test takes into account the magnitude of the differences in substitution rate, so that those comparisons with a greater difference in rate between endosymbionts and nonendosymbionts were more heavily weighted.

Likelihood Ratio Tests

For each comparison, the hypothesis of a difference in substitution rate between the endosymbiotic and nonendosymbiotic clades was tested using a likelihood ratio test (Felsenstein 1981; Huelsenbeck and Rannala 1997). Using Rhino (Rhino: phylogenetic models for estimating rates and dates, version 1.02, available at http://evolve.zoo.ox.ac.uk/software/Rhino/main.html), substitution rates were estimated under two models, both using an HKY + Γ model of substitution (see fig. 2). In the first model, an average substitution rate for all taxa in the comparison was estimated, whereas in the second model, one substitution rate was estimated for the endosymbiont clade and another for the other species in the comparison. To test whether the two rates were significantly different, twice the difference of the log likelihoods of the two models was compared to a chi-squared distribution with one degree of freedom. The substitution rates estimated under the second model were then compared to determine whether the endosymbiont rate was higher than that of the nonendosymbionts and outgroup.

Relative Rate Test

The relative rate test uses distance methods to estimate the difference in the number of substitutions between two closely related taxa in comparison with an outgroup species (Sarich and Wilson 1973). For this analysis, one endosymbiont and one nonendosymbiont were chosen randomly from each independent comparison, and the same outgroups were used as for the Rhino analysis. Branch lengths for the two target species were estimated using an HKY distance matrix generated by PAUP*, and the ratio of the branch lengths was compared to determine whether the endosymbiotic species had a longer branch, indicating a higher substitution rate, than the nonendosymbiotic species.

A + T Content of Endosymbiont 16S Sequences

To test whether endosymbiont 16S genes tend to have higher A + T contents than those of their free-living relatives, the sequence data and comparisons used in the Rhino analysis (see table 1) were used. For each independent comparison, the average A + T content of the 16S gene was calculated for the endosymbiont clade and the nonendosymbiont clade, and the values were compared to determine which was higher. A signed-ranks test was performed across all comparisons (table 3).

A + T Content of Bacterial Whole Genomes

The second data set analyzed comprised all bacterial whole genomes available from GenBank. The average A + T content of the genome was calculated for each of these 48 species. 16S genes from each species were aligned, and maximum likelihood and LogDet phylogenetic trees were constructed, as described above. Species were designated as host-dependent if they had never been recorded as growing outside a host in nature. Independent comparisons of host-dependent and non-host-dependent species pairs were chosen if they were supported by both ML and LogDet trees. For each of these seven comparisons (table 4), the A + T contents of the two species were

Table 1
Independent Comparisons of Endosymbiotic and Nonendosymbiotic Taxa, Used in the Rhino (All Taxa) and Relative Rate Tests (Species Marked with an Asterisk)

Taxonomic Group	Endosymbiont Species (listed by host species name)	Nonendosymbiont Species	Outgroup
α-Proteobacteria	**Oligochaetes:** *Olavius loisae**	*Azospirillum* sp.*, Uncultured Arctic bacterium 95B2	*Novosphingobium subarcticum**
β-Proteobacteria 1	**Mealybugs:** *Antonina crawii*, *Dysmicoccus neobrevipes*, *Planococcus citri**, *Pseudococcus longispinus*	*Burkholderia pseudomallei**, *B. cepacia*, *B. ambifaria*	Soil bacterium SCI84*
β-Proteobacteria 2	**Trypanosomes:** *Blastocrithidia culicis**, *Crythidia* sp.	*Achromobacter xylosoxidans*, Arsenite oxidizing bacterium NT10, *Bordetella trematum**, *B. avium*, Uncultured bacterium GIF8	Uncultured bacterium MS8*
δ-Proteobacteria	**Oligochaetes:** *Olavius algarvensis**	*Desulfobacterium indolicum**, *Desulfonema magnum*, *Desulfosarcina variabilis*, Uncultured bacterium Eel-36e1H6	*Polyangium cellulosum**
Flavobacteria	**Cockroaches:** *Blaberus craniifer*, *Blatella germanica*, *Cryptocercus clevelandi*, *C. darwini**, *C. garciai*, *C. punctulatus*, *C. relictus*, *C. wrighti*, *Nauphoeta cinerea*, *Periplaneta americana*, *P. australasiae*, *Pycnoscelus surinamensis*, *Mastotermes darwiniensis*	*Chryseobacterium meningosepticum*, *Cytophaga* sp., Flavobacteriaceae bacterium, *Flexibacter echinicida*, *Ornithobacterium rhinotracheale*, *Polaribacter* sp.*, Uncultured Cytophagales bacterium	Uncultured Bacteroidetes*
γ-Proteobacteria 1	**Psyllids:** *Acizzia uncatoides*, *Anomoneura mori*, *Aphalaroida inermis*, *Aphalara longicaudata*, *Arytaina genistae*, *Bactericera cockerelli*, *Boreioglycaspis melaleucae*, *Blastopsylla occidentalis*, *Cacopsylla brunneipennis*, *C. myrthi*, *C. pyri*, *C. peregrina*, *Calophya schini*, *Cecidotrioza sozanica*, *Ctenarytaina eucalypti*, *C. longicauda*, *C. spatulata*, *Diaphorina citri*, *Glycaspis brimblecombei*, *Heteropsylla cubana*, *H. texana*, *Neotriozella hirsuta*, *Pachypsylla celtidis*, *P. celtidismamma*, *P. pallida*, *P. venusta*, *Panisopelma fulvescens*, *Paratrioza cockerelli*, *Psylla buxi*, *P. floccosa*, *P. pyricola*, *Russelliana intermedia*, *Spanioneura fonscolombii*, *Tainarys sordida*, *Trioza eugeniae*, *T. magnoliae*, *T. urticae* **Whiteflies:** *Aleurodicus dugesii*, *Bemisia argentifolii*, *B. tabaci*, *Siphoninus phillyreae**, *Trialeurodes vaporariorum*	*Chromohalobacter marismortui*, *Halomonas marina*, *H. venusta*, *H. campisalis*, *H. salina*, *H. desiderata*, *H. pantelleriensis*, *Zymobacter palmae**	*Marinobacterium georgiense**
γ-Proteobacteria 2	**Shipworms:** *Bankia setacea**, *Teredo bartschi*	Uncultured Arctic bacterium BD213*, Uncultured Arctic bacterium C1B045, Uncultured Arctic bacterium 96A14	*Pseudomonas fragi**
γ-Proteobacteria 3	**Clams:** *Calyptogena elongata*, *C. kilmeri*, *C. magnifica*, *C. phaseoliformis*, *C. pacifica*, *C.* sp.*, *Ectenagena extenta*, *Vesicomya chordata*, *V. gigas*, *V. lepta*	Uncultured Arctic bacterium 96BD19*, Uncultured marine bacterium ZD0405	*Francisella philomiragia**
γ-Proteobacteria 4	**Ticks***: Haemaphysalis longcornis**, *Rhipicephalus sanguineus*	*Coxiella burnetii**, *Legionella parisiensis*	Uncultured bacterium C1B045*
γ-Proteobacteria 5	**Aphids:** *Acyrthosiphon pisum*, *Baizongia pistaciae*, *Cinara tujafilina*, *Chaitophorus viminalis**, *Diuraphis noxia*, *Eriosoma lanuginosum*, *Geoica urticularia*, *Hoplocallis pictus*, *Lachnus roboris*, *Maculolachnus submacula*, *Mindarus victoria*, *Melaphis rhois*, *Myzus persicae*, *Panaphis juglandis*, *Pemphigus betae*, *P. populi*, *Pterocomma populeum*, *Rhopalosiphum maidis*, *R. padi*, *Schizaphis graminum*, *Schlechtendalia chinensis*, *Stomaphis quercus*, *Tetraneura caerulescens*, *Thelaxes suberi*, *Tuberolachnus salignus*, *Uroleucon sonchi*, *Yamatocallis hirayamae*, *Y. tokyoensis*; **Carpenter ants:** *Camponotus balzani*, *C. floridanus*, *C. herculeanus*, *C. ligniperdus*, *C. pennsylvanicus*, *C. rufipes*, *C. silvicola*, *C. sericeiventris*, *C. socius* **Tsetse flies:** *Glossina austeni*, *G. brevipalpis*, *G. fuscipes*, *G. morsitans*, *G. palpalis*, *G. tachinoides*; **Weevils:** *Sitophilus zeamais*, *S. oryzae*	*Brennaria salicis*, *B alni**, *Enterobacter intermedius*, *E. nimipressuralis*, *Erwinia amylovora*, *E. carotovora*, *E. psidii*, *Escherichia coli*, *Klebsiella pneumoniae*, *Kluyvera cryocrescens Pantoea agglomerans*, *P. ananatis*, *Pectobacterium chrysanthemi*, *P. cypripedii*, *Proteus vulgaris*, *Rahnella aquatilis*, *Raoultella ornithinolytica*, *R. planticola*, *Salmonella bongori*, *S. typhi*, *S. typhimurium*, *Serratia entomophila*, *S. prtoeamaculans*, *S. rubidae*, *Shigella boydii*, *S. dysenteriae*, *S. flexneri*, *Yersinia enterocolitica*, *Y. frederiksenii*, *Y. kristensenii*, *Y. pseudotuberculosis*	*Burkholderia ambifaria**

Table 1
Continued

Taxonomic Group	Endosymbiont Species (listed by host species name)	Nonendosymbiont Species	Outgroup
Spirochaetes	**Oligochaetes:** *Olavius loisae**	*Spirochaeta alkalica**, *S. asiatica*, *S. bajacaliforniensis*, *S. isovalerica*, *S. smaragdinae*	*Treponema* sp.*
Ascomycota 1	**Anobiid beetles:** *Stegobium paniceum*, *Lasioderma serricorne**	*Aureobasidium pullulans*, *Botryosphaeria ribis*, *Delphinella strobiligena*, *Guignardi endophyllicola*, *Phyllosticta pyrolae**	*Leucostoma persooni**
Ascomycota 2	**Aphids:** *Hamiltonaphis styraci*; **Planthoppers:** *Laodelphax striatellus*, *Nilaparvata lugens*, *Sogatella furcifera**	*Claviceps purpurea*, *C. africana*, *Cordyceps*, *ophioglossoides*, *Hypocrea koningii*, *H. rufa**, *Metarhizium anisopliae*	*Leucostoma persoonii**

compared, and a signed-ranks test was performed across all comparisons.

We then tested whether, for these 48 bacterial species, the base composition of the 16S gene tends to reflect that of the whole genome. This cannot be done by a simple species-by-species comparison of 16S and whole-genome A + T content, as this does not take into account the effects of phylogenetic nonindependence. We therefore chose 15 phylogenetically independent comparisons, each consisting of two species, from the bacterial tree described above (table 5). Pairs were chosen to maximize the total number of comparisons, while excluding pairs of species separated by extremely short branches, such as sister subspecies. For each comparison, the difference between the two species in A + T for the 16S gene (AT_{16S}) was compared to the difference in whole genome A + T (AT_{WG}). If the base composition of 16S reflects that of the whole genome with only random fluctuations, then the median value of $|AT_{WG}| - |AT_{16S}|$ should be zero. A predominance of positive values would indicate that the base composition of the 16S gene tends to differ less than that of the whole genome, whereas more negative values would suggest that the A + T content of the 16S gene differs more.

Table 2
Results of Relative Rate and Rhino (Maximum Likelihood) Tests of Rate Differences Between Endosymbiont and Nonendosymbiont Taxa for the 16S Gene

Comparison	Relative Rate Test	Rhino (likelihood ratio test)
α-Proteobacteria	0.767	0.674
β-Proteobacteria 1	7.864	10.677**
β-Proteobacteria 2	2.707	3.107*
δ-Proteobacteria	0.031	0.220
Flavobacteria	1.221	1.497*
γ-Proteobacteria 1	3.988	5.622**
γ-Proteobacteria 2	0.830	1.322
γ-Proteobacteria 3	2.635	3.119*
γ-Proteobacteria 4	41.81	20.285**
γ-Proteobacteria 5	3.215	5.999**
Spirochaetes	2.273	3.147**
Ascomycota 1	0.594	0.371
Ascomycota 2	1.255	26.491**
Signed-ranks test	$P = 0.05$	$P = 0.014$

NOTE.—The figure reported for the relative rate test is the ratio of the endosymbiont branch length to nonendosymbiont branch length, and for the Rhino test it is the ratio of the endosymbiont rate to the nonendosymbiont rate. Asterisks indicate significance, * $P = 0.05$, ** $P = 0.0005$.

Results

Phylogenetic Analysis

The estimated phylogenetic tree for the bacterial 16S sequences is shown in figure 1. This tree is in broad agreement with previous phylogenies, although it may differ in details of topology with other published trees (e.g., the placement of weevil endosymbionts in a clade with aphid endosymbionts, cf. Heddi et al. 1998). A phylogenetic analysis, using the methods above, of 17 23S sequences from endosymbionts of aphids, psyllids,

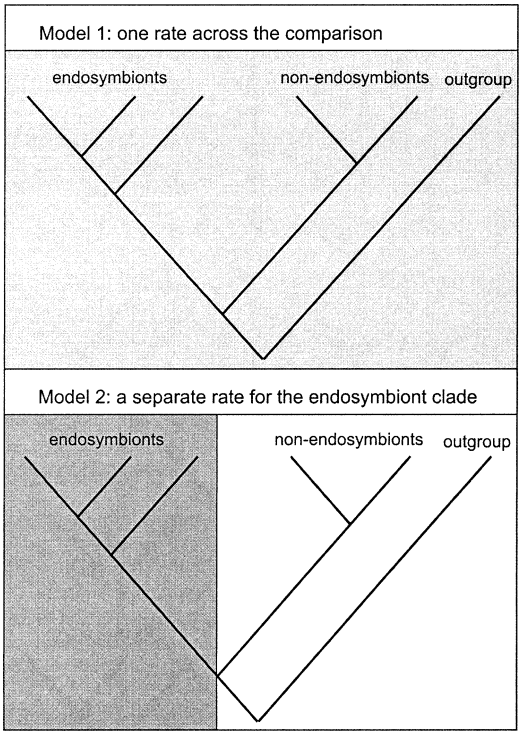

FIG. 2.—A comparison between a monophyletic group of endosymbionts, a closely related group of nonendosymbionts, and an outgroup. For Rhino analyses, the substitution rate is estimated under two models: first with a single HKY + Γ rate across the whole comparison, and then with separate rate estimates for the endosymbiotic clade and the rest of the comparison. The test statistic is twice the difference in the log likelihood values of the two models. The two rates estimated under the second model can be compared to determine whether endosymbionts have a higher rate of evolution than the rest of the comparison.

whiteflies, and tsetse flies and other γ-proteobacterial nonendosymbiotic species, confirmed the relationships of these species (results not shown). For the purposes of this analysis, the only critical feature of the topology is the phylogenetic independence of the comparisons listed in table 1. If the phylogenetic independence of these comparisons is shown to be compromised by future phylogenetic studies, the results of this analysis may need to be reexamined. However, we have chosen them based on the phylogenetic information available to us.

Substitution Rate

Both Rhino and relative rate tests showed that endosymbionts tend to have significantly faster rates of substitution than their free-living relatives (Rhino: 10/13 comparisons, $P = 0.014$; relative rate: 9/13 comparisons, $P = 0.05$; table 2). Likelihood ratio tests indicated that, of the 10 positive comparisons in the Rhino analysis, nine had significantly higher rates of evolution for the endosymbiotic clade than the rest of the species in the comparison. None of the three negative comparisons were significant. Results were consistent between Rhino and relative rate tests for all but one comparison (γ-Proteobacteria 2). For this comparison, a nonsignificant increase in substitution rate in the endosymbiont clade was detected in the Rhino analysis, while in the relative rate test the nonendosymbiotic species had a very marginally higher rate.

The magnitude of the rate difference between endosymbionts and nonendosymbionts estimated using Rhino and relative rate tests was similar for 11 of the comparisons. For the two others (δ-Proteobacteria and Ascomycota 2), the ratio of the rates estimated using Rhino was approximately an order of magnitude higher than that using the relative rate test. There are several differences between the methods, which may account for this variation, such as the method of estimation of branch lengths, inclusion of the outgroup in the Rhino estimation of nonendosymbiont rates, and the number of taxa used (the relative rate test is performed on a single branch from each clade, and therefore utilizes only a small, possibly nonrepresentative sample of the data available). However, in each case the direction of the rate differences is the same for both methods, and the results of both statistical analyses are congruent.

A + T Content

Endosymbiont 16S genes were not significantly more A + T rich than those of nonendosymbionts (signed-ranks test: 9/13 comparisons, $P = 0.076$; table 3). However, our analysis of whole bacterial genomes using phylogenetically independent comparisons showed that host-dependent bacteria (both symbiotic and parasitic) tend to have higher genomic A + T content than non-host-dependent bacteria (signed-ranks test: 6/7 comparisons, $P = 0.0139$; table 4). The difference in base composition between host-dependent and non-host-dependent bacteria is smallest for the one negative comparison (*Corynebacterium glutamicum* versus *Mycobacterium leprae*).

Table 3
Difference in Average 16S A + T Content Between Endosymbiont and Nonendosymbiont Clades

	16S A + T Content (%)	
Comparison	Endosymbiont Clade	Nonendosymbiont Clade
α-Proteobacteria	42.62	44.25
β-Proteobacteria 1	44.09	44.55
β-Proteobacteria 2	47.23	44.54
δ-Proteobacteria	45.80	46.91
Flavobacteria	53.59	49.82
γ-Proteobacteria 1	62.85	43.85
γ-Proteobacteria 2	46.60	46.21
γ-Proteobacteria 3	50.34	49.05
γ-Proteobacteria 4	47.38	46.25
γ-Proteobacteria 5	50.75	45.38
Spirochaetes	45.88	45.23
Ascomycota 1	52.40	52.04
Ascomycota 2	49.48	51.81
Signed-ranks test	$P = 0.076$	

We tested whether the different results for the 16S and whole genome data sets could be due to constraint on A + T content in the 16S gene, by comparing the change in A + T in the 16S and the whole genome between pairs of bacterial species. For all 15 comparisons, the difference in A + T content between the pairs of species was greater across the whole genome than in the 16S gene (sign test $P = 0.00003$; table 5), indicating that the base composition of the 16S gene differs in a biased manner from that of the whole genome.

Discussion

Rate Acceleration in Endosymbionts

Our analysis of 13 independent comparisons demonstrates that endosymbiotic microorganisms from a wide taxonomic and ecological range exhibit a significantly faster rate of substitution than their nonendosymbiotic relatives, a finding consistent with earlier results for a much smaller sample of endosymbiont sequences (Moran 1996; Thao et al. 2000). Although our results are based on a single gene, 16S rRNA, which is the only gene available for the majority of endosymbiont taxa, similar results have been reported for many genes, of widely varying function, in *Buchnera*, the endosymbionts of aphids (Moran 1996; Brynnel et al. 1998; Clark, Moran, and Baumann 1999; Shigenobu et al. 2001; Wernegreen, Richardson, and Moran 2001).

Our results are also consistent with Ohta's (1972) prediction that decreased N_e will lead to an increase in substitution rate. There are, however, several other possible explanations for this increase in rate. A change in selective regime, either increasing the strength of positive selection or relaxing selective constraint, could increase the rate of fixation of mutations. Positive selection leading to increased fixation of nonsynonymous mutations in genes involved in host interaction, such as those coding for outer membrane proteins, has been demonstrated for some parasites (Haydon et al. 2001; Urwin et al. 2002). However, this is less likely to occur in beneficial symbioses, in which conflict between host and symbiont

Table 4
Comparisons of Total Genomic A + T Content in Host-Dependent and Non-Host-Dependent Bacteria

Comparisons		Genomic A + T Content (%)	
Non-Host-Dependent	Host-Dependent	Non-Host-Dependent	Host-Dependent
Synechocystis sp.	*Rickettsia conorii*	52.28	67.56
Pseudomonas aeruginosa	*Neisseria meningitidis*	33.44	48.47
Thermotoga maritima	*Helicobacter pylori*	53.75	61.13
Vibrio cholerae	*Pasteurella multocida*	52.70	59.60
Salmonella typhimurium	*Yersinia pestis*	47.78	52.39
Clostridium acetobutylicum	*Mycoplasma pulmonis*	69.07	73.36
Corynebacterium glutamicum	*Mycobacterium leprae*	46.19	42.22
	Signed-ranks test	$P = 0.014$	

is expected to be reduced (Jiggins, Hurst, and Yang 2002), and in fact *Buchnera* has lost several genes required for biosynthesis of cell-surface components involved in cell defense, suggesting that it does not need to defend itself against the host immune system (Shigenobu et al. 2000).

It is more plausible that endosymbionts might experience a relaxation of purifying selection across loci, as their environment is in some ways more stable than that of free-living microorganisms. It is not possible to distinguish between the effects of decreased selective constraints and reduction in N_e based on the 16S data presented here. However, these two processes are expected to affect intraspecific variation in different ways—polymorphism will increase under relaxed selection, but decrease with a reduction in N_e. A study designed to investigate this question by comparing intraspecific variation in three genes found that polymorphism was significantly lower in *Buchnera* than in enteric bacteria (Funk, Wernegreen, and Moran 2001), suggesting that, at least for this endosymbiont, rate acceleration is more likely to be the result of changes in N_e than reduced constraint.

Another possible cause of rate acceleration is an increase in the mutation rate in endosymbionts, as suggested by Itoh, Martin, and Nei (2002). Such a change in mutation rate has been hypothesized for several lineages of mutualistic fungi, due to an increase in desiccation and UV-exposure associated with moving from a subterranean niche while free-living to growing on the substrate once in symbiosis with an alga or liverwort (Lutzoni and Pagel 1997). It is not clear whether the many lineages of endosymbionts considered here have experienced a similar common ecological change which could trigger an increase in the rate of mutation, as various species of endosymbionts are housed intracellularly and intercellularly in different kinds of organs and many parts of the host body, in a taxonomically diverse range of host species.

The effects of increased mutation rate can be distinguished from those of reduced N_e by examining the types of mutations which are being fixed in the population. An increase in overall mutation rate, without any concomitant change in selective regime, will lead to an increase in the fixation of both synonymous and nonsynonymous mutations. A decrease in N_e, however, will result in synonymous mutations continuing to fix at the same rate, whereas a larger proportion of nonsynonymous mutations will be under the selection threshold and will drift to fixation. The ratio of synonymous to

Table 5
Differences in Genomic and 16S A + T Content for Independent Comparisons of Bacterial Species

Independent Comparisons		Difference in Whole Genome A + T %	Difference in 16S A + T %
Yersinia pestis	*Salmonella typhimurium*	0.046	0.003
Haemophilus influenzae	*Vibrio cholerae*	0.096	0.018
Xylella fastidiosa	*Pseudomonas aeruginosa*	0.139	0.003
Neisseria meningitidis	*Ralstonia solanacearum*	0.155	0.000
Brucella melitensis	*Agrobacterium tumefaciens*	0.022	0.011
Rickettsia prowazekii	*Caulobacter vibrioides*	0.382	0.050
Nostoc sp.	*Synechocystis* sp.	0.061	0.002
Corynebacterium glutamicum	*Mycobacterium leprae*	0.040	0.000
Aquifex aeolicus	*Thermotoga maritima*	0.028	0.015
Campylobacter jejuni	*Deinococcus radiodurans*	0.365	0.004
Fusobacterium nucleatum	*Chlamydia pneumoniae*	0.134	0.011
Borrelia burgdorferi	*Treponema pallidum*	0.242	0.086
Ureaplasma urealyticum	*Mycoplasma pulmonis*	0.012	0.011
Clostridium acetobutylicum	*Streptococcus pneumoniae*	0.088	0.007
Staphylococcus aureus	*Listeria innocua*	0.046	0.028
	Sign test	$P = 0.00003$	

NOTE.—Differences are given as absolute values.

nonsynonymous substitutions is therefore expected to remain constant under the first hypothesis, and increase under the second. Estimating separate synonymous and nonsynonymous substitution rates is not possible for 16S sequences, but several studies of protein-coding genes in *Buchnera* indicate that an increase in this ratio is found across a wide range of genes in this endosymbiont (Moran 1996; Brynnel et al. 1998; Clark, Moran, and Baumann 1999; Wernegreen, Richardson, and Moran 2001; Moya et al. 2002). Once protein coding genes are sequenced from a wider range of endosymbiont taxa, it will be possible to perform similar analyses to determine how general these results are. However, considering the data currently available, it seems likely that the increased substitution rates in endosymbiont 16S genes observed here are due to a reduction in effective population size.

A + T Content of Endosymbionts and Other Host-Dependent Bacteria

The A + T-rich genomes of *Buchnera* and *Wigglesworthia* (the endosymbiont of tsetse flies) have been adduced as further evidence of genetic drift resulting from reduced N_e in endosymbiont species. The genomic A + T content of these species ranges from 74% to 78% (Akman et al. 2002; van Ham et al. 2003), and this A + T enrichment is thought to be the result of the increased dominance of mutational bias over selection for translational efficiency and preservation of gene function (Clark, Moran, and Baumann 1999; Shigenobu et al. 2001; Palacios and Wernegreen 2002). This pattern of A + T-biased base composition has been considered to be a common property of vertically transmitted endosymbiont lineages (Moran and Baumann 2000). However, a significant increase in A + T content in endosymbiont 16S sequences is not observed in our analysis.

A probable explanation for the absence of significant base compositional change is that selection for function and rRNA secondary structure stability is overcoming mutational biases in these sequences (Muto and Osawa 1987; Sueoka 1988). Although endosymbiont 16S genes have accumulated a number of slightly deleterious mutations which have destabilized their rRNA secondary structure (Lambert and Moran 1998), this degeneration may be held in check to some extent by selection against increased A + T counteracting the mutational bias toward A + T. This interpretation is consistent with the results of our analysis comparing A+ T content in 16S genes and whole genomes, for which we found that the base composition of 16S genes varies significantly less than that of the rest of the genome for a wide range of bacteria (see table 5). This suggests that the rate acceleration observed for endosymbionts in this analysis would be more pronounced at other loci, as has been observed for *Buchnera* (Clark, Moran, and Baumann 1999).

Further sequence data will be required to determine whether the pattern of A + T-biased base composition observed in *Buchnera* and *Wigglesworthia* genomes, and in several genes in *Candidatus* Carsonella, is a general feature of endosymbiont genomes. There are a number of plausible causes of such a bias (reviewed in Moran 2002), including biased nucleotide pools favoring A and T, as suggested by Rocha and Danchin (2002). This hypothesis requires further examination, as host-dependent bacteria vary both in their nucleotide biosynthesis ability and in their location within the host, and so would not be expected to display a consistent base compositional response to host-dependence (Moran 2002). Our comparative analysis of bacterial whole genomes found, however, that host-dependent bacteria have significantly increased A + T content, which is consistent with this hypothesis. Although only seven independent comparisons were analyzed, the results suggest that the observed association between host-dependence and base composition is not an artifact of phylogenetic structure. More genomic data would be needed to confirm this pattern, and to examine the mechanism involved.

Acknowledgments

We thank the members of the Centre for the Study of Evolution at the University of Sussex, in particular Adam Eyre-Walker, John Maynard Smith, and Noel Smith, for valuable discussion and comments on the manuscript.

Literature Cited

Akman, L., A. Yamashita, H. Watanabe, K. Oshima, T. Shiba, M. Hattori, and S. Aksoy. 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. Nature Genet. **32**:402–407.

Bromham, L., M. Woolfit, M. S. Y. Lee, and A. Rambaut. 2002. Testing the relationship between morphological and molecular rates of change along phylogenies. Evolution **56**:1921–1930.

Brynnel, E. U., C. G. Kurland, N. A. Moran, and S. G. E. Andersson. 1998. Evolutionary rates for *tuf* genes in endosymbionts of aphids. Mol. Biol. Evol. **15**:574–582.

Buchner, P. 1965. Endosymbiosis of animals with plant microorganisms. Interscience Publishers, New York.

Clark, M. A., L. Baumann, M. A. Munson, P. Baumann, B. C. Campbell, J. E. Duffus, L. S. Osborne, and N. A. Moran. 1992. The eubacterial endosymbionts of whiteflies (Homoptera: Aleyrodoidea) constitute a lineage distinct from the endosymbionts of aphids and mealybugs. Curr. Microbiol. **25**:119–123.

Clark, M. A., N. A. Moran, and P. Baumann. 1999. Sequence evolution in bacterial endosymbionts having extreme base compositions. Mol. Biol. Evol. **16**:1586–1598.

DeSalle, R., and A. R. Templeton. 1988. Founder effects and the rate of mitochondrial DNA evolution in Hawaiian Drosophila. Evolution **42**:1076–1084.

Easteal, S., and C. Collet. 1994. Consistent variation in amino acid substitution rate, despite uniformity of mutation rate—protein evolution in mammals is not neutral. Mol. Biol. Evol. **11**:643–647.

Eyre-Walker, A., P. D. Keightley, N. G. C. Smith, and D. Gaffney. 2002. Quantifying the slightly deleterious mutation model of molecular evolution. Mol. Biol. Evol. **19**:2142–2149.

Felsenstein, J. 1974. The evolutionary advantage of recombination. Genetics **78**:737–756.

———. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. **17**:368–376.

———. 1985. Phylogenies and the comparative method. Am. Nat. **125**:1–15.

Funk, D. J., J. J. Wernegreen, and N. A. Moran. 2001. Intraspecific variation in symbiont genomes: bottlenecks and the aphid–buchnera association. Genetics **157**:477–489.

Gillespie, J. H. 1994. Substitution processes in molecular evolution III: deleterious alleles. Genetics **138**:943–952.

Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford University Press, Oxford.

Harvey, P. H., and A. Rambaut. 1998. Phylogenetic extinction rates and comparative methodology. Proc. R. Soc. Lond. Ser. B. **265**:1691–1696.

Hasegawa, M., H. Kishino, and T. A. Yano. 1985. Dating of the human ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22**:160–174.

Haydon, D. T., A. D. Bastos, N. J. Knowles, and A. R. Samuel. 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. Genetics **157**:7–15.

Heddi, A., H. Charles, C. Khatchadourian, G. Bonnot, and P. Nardon. 1998. Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G + C content of an endocytobiotic DNA. J. Mol. Evol. **47**: 52–61.

Huelsenbeck, J. P., and B. Rannala. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science **276**:227–232.

Itoh, T., W. Martin, and M. Nei. 2002. Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. Proc. Natl. Acad. Sci. USA **99**:12944–12948.

Jiggins, F. M., G. D. D. Hurst, and Z. H. Yang. 2002. Host-symbiont conflicts: positive selection on an outer membrane protein of parasitic but not mutualistic Rickettsiaceae. Mol. Biol. Evol. **19**:1341–1349.

Johnson, K. P., and J. Seger. 2001. Elevated rates of non-synonymous substitution in island birds. Mol. Biol. Evol. **18**:874–881.

Kimura, M., and T. Ohta. 1971. On the rate of molecular evolution. J. Mol. Evol. **1**:1–17.

Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, and J. Hey. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics **156**: 1913–1931.

Lambert, J. D., and N. A. Moran. 1998. Deleterious mutations destabilize ribosomal RNA in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA **95**:4458–4462.

Lockhart, P. J., M. A. Steel, M. D. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

Lutzoni, F., and M. Pagel. 1997. Accelerated evolution as a consequence of transitions to mutualism. Proc. Natl. Acad. Sci. USA **94**:11422–11427.

Lynch, M., and J. L. Blanchard. 1998. Deleterious mutation accumulation in organelle genomes. Genetica **103**:29–39.

Martins, E. P., and T. Garland. 1991. Phylogenetic analyses of the correlated evolution of continuous characters—a simulation study. Evolution **45**:534–557.

Mira, A., and N. A. Moran. 2002. Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. Microb. Ecol. **44**:137–143.

Mooers, A. Ø., and E. C. Holmes. 2000. The evolution of base composition and phylogenetic inference. Trends Ecol. Evol. **15**:365–369.

Moran, N. A. 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA **93**:2873–2878.

———. 2002. Microbial minimalism: Genome reduction in bacterial pathogens. Cell **108**:583–586.

Moran, N. A., and P. Baumann. 2000. Bacterial endosymbionts in animals. Curr. Opin. Microbiol. **3**:270–275.

Moya, A. S., A. Latorre, B. Sabater-Munoz, and F. J. Silva. 2002. Comparative molecular evolution of primary (*Buchnera*) and secondary symbionts of aphids based on two protein-coding genes. J. Mol. Evol. **55**:127–137.

Munson, M. A., P. Baumann, and N. A. Moran. 1992. Phylogenetic relationships of the endosymbionts of mealybugs (Homoptera: Pseudococcidae) based on 16S rDNA sequences. Mol. Phylogenet. Evol. **1**:26–30.

Muto, A., and S. Osawa. 1987. The guanine and cytosine content of genomic DNA and bacterial evolution. Proc. Natl. Acad. Sci. USA **84**:166–169.

Ohta, T. 1972. Population size and the rate of evolution. J. Mol. Evol. **1**:305–314.

———. 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. **23**:263–286.

Palacios, C., and J. J. Wernegreen. 2002. A strong effect of AT mutational bias on amino acid usage in Buchnera is mitigated at high-expression genes. Mol. Biol. Evol. **19**:1575–1584.

Pamilo, P., M. Nei, and W. H. Li. 1987. Accumulation of mutations in sexual and asexual populations. Genet. Res. **49**:135–146.

Rambaut, A. 2002. Se-Al sequence alignment editor. Version 2.0. Available at http://evolve.zoo.ox.ac.uk/software/Se-Al/main.html.

Rispe, C., and N. A. Moran. 2000. Accumulation of deleterious mutations in endosymbionts: Muller's ratchet with two levels of selection. Am. Nat. **156**:425–441.

Rocha, E. P. C., and A. Danchin. 2002. Base composition bias might result from competition for metabolic resources. Trends Genet. **18**:291–294.

Sanderson, M. J. 1990. Estimating rates of speciation and evolution—a bias due to homoplasy. Cladistics **6**:387–391.

Sarich, V. M., and A. C. Wilson. 1973. Generation time and genomic evolution in primates. Science **179**:1144–1147.

Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp APS. Nature **407**:81–86.

Shigenobu, S., H. Watanabe, Y. Sakaki, and H. Ishikawa. 2001. Accumulation of species-specific amino acid replacements that cause loss of particular protein functions in Buchnera, an endocellular bacterial symbiont. J. Mol. Evol. **53**:377–386.

Spaulding, A. W., and C. D. von Dohlen. 1998. Phylogenetic characterization and molecular evolution of bacterial endosymbionts in psyllids (Hemiptera : Sternorrhyncha). Mol. Biol. Evol. **15**:1506–1513.

———. 2001. Psyllid endosymbionts exhibit patterns of co-speciation with hosts and destabilizing substitutions in ribosomal RNA. Insect Mol. Biol. **10**:57–67.

Spradling, T. A., M. S. Hafner, and J. W. Demastes. 2001. Differences in rate of cytochrome-*b* evolution among species of rodents. J. Mammal. **82**:65–80.

Sueoka, N. 1988. Directional mutation pressure and neutral molecular evolution. Proc. Natl. Acad. Sci. USA **85**:2653–2657.

Swofford, D. L. 2002. PAUP*. phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Mass.

Tamas, I., L. Klasson, B. Canback, A. K. Naslund, A. S. Eriksson, J. J. Wernegreen, J. P. Sandstrom, N. A. Moran, and S. G. E. Andersson. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science **296**:2376–2379.

Thao, M. L., N. A. Moran, P. Abbot, E. B. Brennan, D. H. Burckhardt, and P. Baumann. 2000. Cospeciation of psyllids and their primary prokaryotic endosymbionts. Appl. Environ. Microbiol. **66**:2898–2905.

Urwin, R., E. C. Holmes, A. J. Fox, J. P. Derrick, and M. C. J. Maiden. 2002. Phylogenetic evidence for frequent positive selection and recombination in the meningococcal surface antigen PorB. Mol. Biol. Evol. **19**:1686–1694.

van Ham, R. C. J. H., J. Kamerbeek, C. Palacios et al. (16 co-authors). 2003. Reductive genome evolution in *Buchnera aphidicola*. Proc. Natl. Acad. Sci. USA **100**:581–586.

Weinreich, D. M. 2001. The rates of molecular evolution in rodent and primate mitochondrial DNA. J. Mol. Evol. **52**:40–50.

Wernegreen, J. J., and N. A. Moran. 1999. Evidence for genetic drift in endosymbionts (*Buchnera*): analyses of protein-coding genes. Mol. Biol. Evol. **16**:83–97.

Wernegreen, J. J., A. O. Richardson, and N. A. Moran. 2001. Parallel acceleration of evolutionary rates in symbiont genes underlying host nutrition. Mol. Phylogenet. Evol. **19**:479–485.

Wu, C. I., and W. H. Li. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA **82**:1741–1745.

Yang, Z. H. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites—approximate methods. J. Mol. Evol. **39**:306–314.

Richard Thomas, Associate Editor

Accepted May 11, 2003