Coevolution between a cockroach and its bacterial endosymbiont: a biogeographical perspective

Jeffrey W. Clark, Shaon Hossain†, Craig A. Burnside‡ and Srinivas Kambhampati*

Department of Entomology, Kansas State University, Manhattan, KS 66506, USA

Cryptocercus are subsocial, xylophagous cockroaches that live in temperate forests. Like other cockroaches, *Cryptocercus* harbour endosymbiotic bacteria in their fat bodies. Two species of *Cryptocercus* occur in the palaearctic, one each in eastern Russia and south-central China. In the USA, there are five species: one in the north-west and four in the south-east. Little is known about the relationship between the Eurasian and North American *Cryptocercus* or the causes of the disjunct distribution. Here, a molecular phylogeny for six out of the seven *Cryptocercus* species and their endosymbionts is inferred in an attempt to understand the evolution and biogeography of the genus. Our analysis showed that the North American *Cryptocercus* are monophyletic, suggesting that a single colonization event was followed by vicariance. There was complete concordance between the host and endosymbiont phylogenetic trees. Divergence estimates based on endosymbiont DNA sequences suggested that the palaearctic and nearctic *Cryptocercus* diverged 70–115 million years (Myr) ago and the eastern- and western-USA species diverged 53–88 Myr ago. These divergence estimates were correlated with biogeographical events, and a hypothesis is presented to explain the current distribution of *Cryptocercus*. Our findings suggest that *Cryptocercus* has had a long evolutionary history, dating back to the Jurassic.

Keywords: *Cryptocercus*; *Blattabacterium*; biogeography, phylogenetic tree

1. INTRODUCTION

Members of the cockroach genus *Cryptocercus* are subsocial, xylophagous insects that live in temperate forests (Cleveland *et al.* 1934). Two species of *Cryptocercus* occur in the palaearctic, one each in eastern Russia (*Cryptocercus relictus*) and south-central China (*Cryptocercus primarius*). Recent work has established the existence of five species in the nearctic (Kambhampati *et al.* 1996; Nalepa *et al.* 1997; Burnside *et al.* 1999). One species, *Cryptocercus clevelandi*, is found in the northwestern USA while the other four, *Cryptocercus darwini*, *Cryptocercus garciai*, *Cryptocercus punctulatus* and *Cryptocercus wrighti*, are found in the southeastern USA.

A prominent feature of the geographical distribution of *Cryptocercus* is the spectacular disjunction: in the USA, the eastern and western species are separated by over 4500 km, the species in the palaearctic are separated by ca. 3500 km, and the palaearctic and nearctic species are separated by over 7000 km. The underlying cause(s) and the temporal dynamics of this disjunct distribution have been the subject of spirited discussion (Grandcolas 1999a,b; Nalepa & Bandi 1999). However, the statements of Grandcolas (1999a) and Nalepa & Bandi (1999) concerning the evolutionary dynamics and present distribution of *Cryptocercus* were merely speculative because reliable estimates of divergence times among *Cryptocercus* species were lacking.

Bacterial endosymbionts, found in many insect orders (Sauer *et al.* 2000; Schroder *et al.* 1996; Dasch *et al.* 1984; O'Neill *et al.* 1992, 1997), provide a potential avenue for estimating divergence times among taxa. Cockroaches,

including *Cryptocercus*, harbour endosymbiotic bacteria within specialized cells (mycetocytes or bacteriocytes) of their fat bodies. These symbionts, *Blattabacterium cuenoti*, infect the developing oocyte before vitellogenesis and are maternally inherited (Sacchi *et al.* 1985). The maternal inheritance and a close physiological relationship with the host make them a useful tool for studies of coevolution. More importantly, the existence of a universal substitution rate (molecular clock) for the bacterial 16S rRNA gene has been demonstrated in several independent studies (e.g. Ochman & Wilson 1987; Moran *et al.* 1993; Bandi *et al.* 1995). The existence of the molecular clock enables estimation of divergence times among the bacterial lineages harboured by the various *Cryptocercus* species. Bandi *et al.* (1995) estimated that the 16S rRNA gene of cockroach endosymbionts evolves at a rate of about 0.6–1% sequence divergence per 50 million years (Myr), a rate similar to that reported for bacteria in general (Ochman & Wilson 1987).

The objective of this study was to investigate the evolution and biogeography of *Cryptocercus*. Specifically, we infer a phylogenetic relationship among the endosymbionts of *Cryptocercus* from the DNA sequence of the 16S rRNA and 23S rRNA genes. We compare the bacterial phylogeny with that of the host, inferred from the DNA sequence of mitochondrial and nuclear rRNA genes, to detect parallel cladogenesis. Previously published molecular-clock values are used to estimate divergence times among *Cryptocercus* species to gain insights into the biogeography of the genus and the probable events that led to the present distribution.

2. MATERIAL AND METHODS

(a) Insects

Cryptocercus darwini, *C. garciai*, *C. punctulatus* and *C. wrighti* were collected in the Appalachian mountains in the eastern USA (for

* Author for correspondence (srini@ksu.edu).

† Present address: Center for Molecular Medicine and Genetics, School of Medicine, Wayne State University, Detroit, MI 48201, USA.

‡ Present address: 100 West College Park Drive, South Georgia College, Douglas, GA 31533, USA.

Table 1. *DNA sequence of the primers used for amplification of genes from the endosymbionts*

(The expected size of the amplified fragment (base pairs) is also shown.)

name	sequence	fragment size
16S rRNA		
33f	TGCAAGTCGAGGGGC	630
660r	TCACCGCTACACCACACATTC	—
660f	GAATGTGTGGTGTAGCGGTGA	640
1294r	GTCGAGTTGCAGACTCCAATC	—
23S rRNA		
23Saf	GACCGATAGTGAACCAGTAC	650
23SAr	CTGCTTCYAAGCCAACMTCC	—
23SBf	CGGGTAAGTTCGACCTGCA	560
23SBr	GAGCCGACATCGAGGTGCCAA	—

details, see Kambhampati *et al.* 1996; Burnside *et al.* 1999). *Cryptocercus cleavelandi* were collected during August 1999 in Oregon. *Cryptocercus relictus* were collected in 1998 near Anisimovka in eastern Siberia.

(b) *DNA extraction, polymerase chain reaction and sequencing*

The endosymbiont 16S rRNA and 23S rRNA gene fragments were amplified from fat-body preparations and the internal transcribed spacer 2 (ITS2) fragment of *Cryptocercus* was amplified from leg muscles as described in Burnside *et al.* (1999). Polymerase chain reaction (PCR) was set up as described in Kambhampati *et al.* (1992). The temperature profile for amplifying the endosymbiont 16S rRNA fragment was 95 °C for 3 min, followed by 40 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s. A final extension step of 72 °C for 7 min was included. For the amplification of the endosymbiont 23S rRNA fragment, an annealing temperature of 45 °C was used for ten cycles followed by 50 °C for 30 cycles; all other parameters were as above.

Primers for the amplification of the host mitochondrial 16S rRNA and 12S rRNA genes were given in Kambhampati *et al.* (1996). Primers for the amplification of the endosymbiont genes (table 1) were designed, based on sequences from GenBank. These primers amplified nearly the entire 16S rRNA gene (*ca.* 1260 bp) and two internal fragments of the 23S rRNA gene totalling *ca.* 1230 bp.

Portions of the host 28S rRNA and 5.8S rRNA genes and the entire ITS2 from the nuclear ribosomal DNA of *Cryptocercus* were amplified using published primers (Campbell *et al.* 1993). The sizes of the fragments ranged from 316 to 359 bp. The amplification conditions were as described for the 23S rRNA gene. All PCR products were gel purified and sequenced as described in Burnside *et al.* (1999). The mitochondrial 16S rRNA and the 12S rRNA gene sequences used here, with the exception of those from *C. relictus*, were previously published in Burnside *et al.* (1999).

For the endosymbiont genes and ITS2, sequences were obtained from multiple individuals collected at the same location to determine the extent of intraspecific variation. Thus, we obtained the DNA sequences from up to five individuals from each of five different locations; the sample sizes for all other locations ranged from one to three individuals. In addition, to detect any variation in the ITS2 region or among bacteria

within a single host, the amplified fragment was cloned into a T-vector (Invitrogen Corporation, Carlsbad, CA, USA) and five randomly chosen clones were sequenced as described in Kambhampati *et al.* (2000). Intra-individual variation was assessed for one specimen from each of the four eastern-USA species.

(c) *Phylogenetic analysis*

The alignment of the 16S rRNA and 12S rRNA genes was that used in Burnside *et al.* (1999) with the addition of sequences from *C. relictus*. The endosymbiont and nuclear sequences were aligned for each specimen using Sequence Navigator (Applied Biosystems, Foster City, CA, USA) or CLUSTALV (Higgins & Sharp 1989). Adjustments to the alignment were made manually as needed. The data were analysed using PAUP 4.0b3a (Swofford 1998); both parsimony and maximum-likelihood methods were used. The heuristic algorithm was used for both models. Gaps were treated as the fifth base. For the maximum-likelihood analysis, base frequencies, between-site variability and rate matrix were estimated during analysis. The Rogers–Swofford method was used to obtain the starting branch lengths. Bootstrapping (10 000 replicates) and decay index (Bremer 1994) were performed to assess branch support. For the phylogenetic analysis of endosymbiont sequences, *Flavobacterium* sp. (GenBank numbers: 16S rRNA, AJ009687 and 23S rRNA, M62807) was used as the outgroup because *B. ceno* is a member of the flavobacteria–bacteroides group (Bandi *et al.* 1994). We used homologous sequences from the locust as the outgroup for phylogenetic analysis of host mitochondrial sequences; ITS2 data for the locust were not available and therefore designated as missing.

The final data set consisted of 16 ingroup taxa. Six species of *Cryptocercus*, five from the USA and one from Russia, were represented. The eastern-USA species were represented by between two and five locations each. The mitochondrial 16S rRNA and 12S rRNA data set, the ITS2 data set and the endosymbiont dataset were all matched exactly with regard to the ingroup taxon sampling. We used COMPONENT Lite (R. Page, University of Glasgow, UK) to compare the trees based on the host and endosymbiont genes. The sequences reported in this paper have been deposited in GenBank under accession numbers AF322458–AF322523.

3. RESULTS

There was no DNA sequence variation among bacteria within a host. Among hosts from the same site the mean bacterial sequence divergence was 0.06% for the 16S rRNA gene and 0.2% for the 23S rRNA gene. Similarly, we detected low ($\leq 1\%$) intra-individual and intra-location sequence variation in ITS2. Similar patterns were reported for the mitochondrial 16S rRNA and 12S rRNA genes (Burnside *et al.* 1999). Intraspecific variation, where present, was low: $\leq 2\%$ for host genes (Burnside *et al.* 1999) and $\leq 0.4\%$ for endosymbiont genes. Therefore, we chose one sequence from each location at random for inclusion in the final data set. The choice of the individual sequence did not affect the topology of the phylogenetic tree.

The aligned data set for the host genes consisted of 1317 characters (895 characters from mitochondrial DNA (mtDNA) and 422 from nuclear rDNA). Out of these, 271 characters were parsimony informative. The endosymbiont data set consisted of 2523 characters, out of which

Figure 1. Phylogenetic trees for *Cryptocercus* (host) and *Blattabacterium cuenoti* (endosymbiont) based on unweighted parsimony analysis. Gaps were included and treated as the fifth base. The host tree (1042 steps, consistency index 0.74, retention index 0.65, rescaled consistency index 0.48) is based on the DNA sequence of two mitochondrial rRNA genes and the nuclear internal transcribed spacer 2 region. The endosymbiont tree (313 steps, consistency index 0.77, retention index 0.65, rescaled consistency index 0.50) is based on the 16S rRNA and 23S rRNA genes; a strict consensus of three equally parsimonious trees is shown. The numbers above the branches are bootstrap values (percentages) and those below the branches are decay indexes (steps). The numbers enclosed in circles are the inferred male diploid chromosome numbers (see Kambhampati *et al.* 1996). Multiple locations are represented for the four eastern-USA species identified by the numbers 1 through 4 as follows: 1, *C. garciai*; 2, *C. darwini*; 3, *C. wrighti*; and 4, *C. punctulatus*. The filled circles on two nodes of the endosymbiont tree denote the only conflicts between the host tree and the endosymbiont tree; statistical tests (see § 3) indicated that the similarity between the two trees is significantly ($p < 0.001$) greater than that expected by chance alone.

87 were parsimony informative. Separate analysis of the endosymbiont and host mitochondrial data sets using free-living flavobacteria and locusts, respectively, as outgroups indicated that *C. relictus* is basal to all *Cryptocercus* species included in this study (see § 4). Therefore, *C. relictus* was designated as the outgroup in the analyses described below.

Details of phylogenetic analysis using the mitochondrial 16S rRNA and 12S rRNA genes have been described elsewhere (Burnside *et al.* 1999). With *C. relictus* as the outgroup, the 16S rRNA and 12S rRNA data resulted in two trees, each of 446 steps, in which *Cryptocercus clevelandi* was basal, followed by *C. wrighti*, *C. punctulatus*, *C. garciai* and *C. darwini*. The ITS2 data, when analysed alone, resulted in eight equally parsimonious trees, each of 574 steps; the topology of a strict consensus tree differed from that of the tree based on mtDNA data in that *C. punctulatus* was more basal than *C. wrighti*. A combined analysis of mtDNA and ITS2 data resulted in a single

tree of 1042 steps in which *C. clevelandi* was basal, followed by *C. punctulatus*, *C. wrighti*, *C. garciai*, and *C. darwini* (figure 1). All the branches were supported in over 50% of the bootstrap replications and were stable in trees several steps longer than the shortest trees.

The results from the analyses of the bacterial 16S rRNA and the 23S rRNA gene fragments individually are not shown here; the topology of the tree based on the 16S rRNA fragment alone was similar to that of the one based on the combined data set (see below). The topology of the tree based on the 23S rRNA gene fragments alone with the gaps excluded (not shown) was similar to those based on the 16S rRNA gene and the combined data set. An analysis of the combined endosymbiont data set resulted in three equally parsimonious trees, each of 313 steps, a strict consensus of which is shown in figure 1. Most branches were supported in over 50% of the bootstrap replications and in trees several steps longer than the most parsimonious ones. The topology of the tree

Figure 2. Maximum-parsimony tree for *Cryptocercus* species. All characters were given equal weight and gaps were treated as the fifth base. The tree shows the results of two separate analyses: first, the DNA sequences of the mitochondrial DNA and rRNA genes and nuclear internal transcribed spacer 2 (ITS2) region of *Cryptocercus* (tree length 1296 steps, consistency index 0.76, retention index 0.63, rescaled consistency index 0.48) and, second, the DNA sequences of the above host genes and the DNA sequences of the endosymbiont's rRNA genes (tree length 1931 steps, consistency index 0.79, retention index 0.63, rescaled consistency index 0.48). The trees from the two data sets were identical in topology except in one respect: in the tree derived from the host genes, *C. darwini* from Buck Creek Trail and De Soto State Park were sisters, whereas in the tree derived from host and endosymbiont sequences, *C. darwini* from De Soto State Park and Cheaha State Park were sisters; thus the relationship between these three samples is shown as a polytomy. In both trees, however, the relationship between these three samples was fully resolved and supported in over 50% of the bootstrap replications. Numbers above the branches are bootstrap values (percentages) from the first and second analyses, respectively; numbers below the branches are decay indexes (steps) from the first and second analyses, respectively. Homologous DNA sequence of *Locusta migratoria* (Flood *et al.* 1997) was used as the outgroup in the first analysis; ITS2 data for the outgroup were designated missing. In the second analysis, the locust sequence and homologous sequences from *Flavobacterium* sp. were used as outgroups. Multiple locations are represented for the four eastern-USA species, which are identified by the numbers 1 through 4 as follows: 1, *C. garciai*; 2, *C. darwini*; 3, *C. wrighti*; and 4, *C. punctulatus*.

based on maximum likelihood was identical to that of the tree based on parsimony. Excluding gaps from the analysis did not affect the topologies of the trees obtained.

The topologies of trees based on the host and endosymbiont genes, in the context of the relationships between species, were identical. There were two minor differences between the host tree and the endosymbiont

tree; both differences were for infraspecific nodes (see figure 1). A test for congruency indicated that the host and endosymbiont data sets had significant similarities between them. The value for the path-difference matrix between the two topologies was 6.1, a value significantly ($p < 0.001$) lower than expected if the similarity between the topologies was due to chance alone.

4. DISCUSSION

Little is known about the evolution of *Cryptocercus* on a global scale. Although *Cryptocercus* occurs in Eurasia and the USA, the vast majority of what is known is based on the North American species and the majority of that knowledge is based on those that occur in the eastern USA. The present study included DNA sequences from six out of the seven known species as well as DNA sequences from the endosymbionts of those six species. Our analyses indicated that there is parallel cladogenesis between the hosts (*Cryptocercus* species) and the endosymbionts (*B. cuenoti*). In other words, there has not been a detectable horizontal transfer of endosymbionts between the hosts. The strong concordance between host and endosymbiont phylogeny enables us to make inferences about the biogeography of the host because the clock-like evolution of the bacterial 16S rRNA gene allows for the estimation of divergence times (Ochman & Wilson 1987; Moran *et al.* 1993; Bandi *et al.* 1995).

Sequence divergence estimates based on the endosymbiont 16S rRNA gene indicated that the endosymbionts of *C. relictus* and the nearctic *Cryptocercus* diverged 70–115 Myr ago (2.8% mean sequence divergence) and those of *C. clevelandi* and the eastern-USA species diverged 53–88 Myr ago (2.1% mean sequence divergence). Nalepa *et al.* (1997) proposed a divergence time of 25–70 Myr for the endosymbionts of *C. clevelandi* and *C. punctulatus*, which differs from our estimate. However, our divergence estimates are based on a larger number of samples and are therefore likely to be closer to the true value. Endosymbionts of *C. punctulatus* and *C. wrighti* diverged during the Oligocene, 23–38 Myr ago, and the divergence between the endosymbionts of *C. darwini* and *C. garciai* occurred during the Miocene, 13–20 Myr ago. The above figures represent divergence times for the various endosymbiont lineages and are therefore the minimum divergence times for the corresponding hosts. In other words, it is reasonable to assume that divergence of the host lineages began several million years before that of the bacterial lineages.

An unresolved issue with regard to *Cryptocercus* is the factors responsible for the aforementioned global disjunct distribution. This is largely due to the fact that there has been no attempt to infer the evolutionary relationship between *C. relictus* and its Nearctic congeners or estimate divergence times between these species. The analyses presented here shed considerable light on this issue. Our results from both the host and the endosymbiont genes (figure 2) indicate that the nearctic *Cryptocercus* species are monophyletic. In addition, a comparison of the phylogenetic trees derived from the host's mitochondrial and nuclear genes and the endosymbiont rRNA genes indicates that they are congruent, i.e. *B. cuenoti* and the host *Cryptocercus* species have diverged concordantly.

Our results lead us to rule out the possibility that the current distribution of nearctic *Cryptocercus* is the result of multiple colonizations over land bridges and subsequent vicariance caused by climatic changes associated with Pleistocene glaciations as proposed by Grandcolas (1999a). Our data suggest that the split between *C. clevelandi* and the species in the eastern USA occurred 53–88 Myr ago, rather than during the Pleistocene glaciations, which occurred a mere 0.5–2.0 Myr ago. Furthermore, monophyly of the nearctic species indicates that they are descended from the same ancestral lineage, rather than from successive invasions.

Based on the phylogeny of *Cryptocercus* species, the phylogeny of their endosymbionts and the estimated divergence times, we offer the following scenario for the current global distribution of *Cryptocercus*. During the Jurassic (144–213 Myr ago) the ancestor of extant *Cryptocercus* inhabited the temperate deciduous forests of the Arcto-Tertiary complex in the extreme northern regions of the Northern Hemisphere. During the Cretaceous (65–144 Myr ago), Asia and North America began to drift apart, but remained intermittently connected by land bridges and eastern and western North America were separated by a shallow inland sea (King 1958). A general cooling trend began during the Cretaceous, which forced the Arcto-Tertiary community (and the ancestor of *Cryptocercus*) to move south into Asia and North America (MacGinitie 1958), resulting in two major lineages. In addition, the ancestral populations of *Cryptocercus* that migrated into North America were split into eastern and western populations because of the prior presence of the inland sea, resulting in two further lineages. By the close of the Palaeocene and the beginning of the Eocene, 58 Myr ago, the inland sea had receded, once again connecting eastern and western North America. However, our divergence estimates indicate that there was no reproductive contact between the eastern and western North American lineages of *Cryptocercus* after the inland sea had receded. The Miocene (5–24 Myr ago) saw further cooling and a drier climate, which led to the formation of extensive grasslands across North America and reinforced the vicariance of plants and animals. Climatic fluctuations of the Pliocene, beginning about 5 Myr ago, became more pronounced during the Pleistocene, further reducing gene flow among the eastern populations. Among the eastern species, *C. wrighti* and *C. punctulatus* are the older lineages whereas *C. darwini* and *C. garciai* are more recent. The speciation events in the eastern USA probably took place during the past 25 Myr. Thus, our results indicate that *Cryptocercus* is not a recent arrival on the evolutionary scene, as suggested by Grandcolas (1999a), rather it belongs to a lineage whose history dates back to the Jurassic. It is not clear what caused the disjunction between *C. relictus* and *C. primarius*; the inclusion of *C. primarius* in a phylogenetic analysis along with other *Cryptocercus* species will be necessary to resolve this issue and fully comprehend the biogeography of the genus. For the scenario proposed above to be plausible, *C. relictus* and *C. primarius* should be monophyletic, representing one of the two major lineages that arose during the migration of *Cryptocercus* to Eurasia and North America.

We thank B. Steinmiller for help in collecting specimens in the western USA and E. Zacharov for providing specimens of *C. relictus*. This study was supported by National Science Foundation grant DEB 9806710 to S.K. This is journal article 00-424-J of the Kansas Agricultural Experiment Station.

REFERENCES

- Bandi, C., Damiani, G., Magrassi, L., Grigolo, A., Fani, R. & Sacchi, L. 1994 Flavobacteria as intracellular symbionts in cockroaches. *Proc. R. Soc. Lond. B* **257**, 43–48.
- Bandi, C., Sironi, M., Damiani, G., Magrassi, L., Nalepa, C. A., Laudani, U. & Sacchi, L. 1995 The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. *Proc. R. Soc. Lond. B* **259**, 293–299.
- Bremer, K. 1994 Branch support and tree stability. *Cladistics* **10**, 295–304.
- Burnside, C. A., Smith, P. T. & Kambhampati, S. 1999 Three new species of the wood roach, *Cryptocercus* (Blattodea: Cryptocercidae), from the eastern United States. *J. Kansas Entomol. Soc.* **72**, 361–378.
- Campbell, B. C., Steffen-Campbell, J. D. & Werren, J. H. 1993 Phylogeny of the *Nasonia* species complex (Hymenoptera: Pteromalidae) inferred from an internal transcribed spacer (ITS2) and 28S rDNA sequences. *Insect Mol. Biol.* **2**, 225–237.
- Cleveland, L., Hall, S. R., Sanders, E. P. & Collier, J. 1934 The wood feeding roach *Cryptocercus*, its protozoa, and the symbiosis between protozoa and roach. *Mem. Am. Acad. Arts Sci.* **17**, 185–342.
- Dasch, G. A., Weiss, E. & Chang, K. P. 1984 Symbionts of insects. *Bergey's Man. Syst. Bacteriol.* **1**, 811–833.
- Flook, P. K., Rowell, C. H. & Gellissen, G. 1997 The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome. *J. Mol. Evol.* **41**, 928–941.
- Grandcolas, P. 1999a Systematics, endosymbiosis and biogeography of *Cryptocercus clevelandi* and *Cryptocercus punctulatus* (Blattaria: Polyphagidae) from North America: a phylogenetic perspective. *Annls Entomol. Soc. Am.* **92**, 285–291.
- Grandcolas, P. 1999b Reconstructing the past of *Cryptocercus* (Blattaria: Polyphagidae): phylogenetic histories and stories. *Annls Entomol. Soc. Am.* **92**, 303–307.
- Higgins, D. M. & Sharp, P. M. 1989 Fast and sensitive multiple sequence alignment on a microcomputer. *Comput. Appl. Biosci.* **5**, 151–153.
- Kambhampati, S., Black IV, W. C. & Rai, K. S. 1992 RAPD of mosquito species and populations: techniques, statistical analysis and applications. *J. Med. Entomol.* **29**, 939–945.
- Kambhampati, S., Luykx, P. & Nalepa, C. A. 1996 Evidence for sibling species in *Cryptocercus punctulatus*, the wood roach, from variation in mitochondrial DNA and karyotype. *Heredity* **76**, 1–12.
- Kambhampati, S., Völkl, W. & Mackauer, M. 2000 Phylogenetic relationship among genera of Aphidiinae (Hymenoptera: Braconidae) based on DNA sequence of the mitochondrial 16S rRNA gene. *Syst. Entomol.* **25**, 437–445.
- King, P. B. 1958 Evolution of modern surface features of western North America. In *Zoogeography* (ed. C. L. Hubbs), pp. 3–60. Washington, DC: The American Association for the Advancement of Science.
- MacGinitie, H. D. 1958 Climate since the Late Cretaceous. In *Zoogeography* (ed. C. L. Hubbs), pp. 61–80. Washington, DC: The American Association for the Advancement of Science.
- Moran, N. A., Munson, M. A., Baumann, P. & Ishikawa, H. 1993 A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc. R. Soc. Lond. B* **253**, 167–171.
- Nalepa, C. A. & Bandi, C. 1999 Phylogenetic status, distribution and biogeography of *Cryptocercus* (Dictyoptera: Cryptocercidae). *Annls Entomol. Soc. Am.* **92**, 292–302.
- Nalepa, C. A., Byers, G. W., Bandi, C. & Sironi, M. 1997 Description of *Cryptocercus clevelandi* (Dictyoptera: Cryptocercidae) from the northwestern United States, molecular analysis of bacterial symbionts in its fat body, and notes on biology, distribution, and biogeography. *Annls Entomol. Soc. Am.* **90**, 416–424.
- Ochman, H. & Wilson, A. C. 1987 Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* **26**, 74–86.
- O'Neill, S. L., Giordano, R., Colbert, A. M. E., Karr, T. L. & Robertson, H. M. 1992 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc. Natl Acad. Sci. USA* **89**, 2699–2702.
- O'Neill, S. L., Hoffmann, A. A. & Werren, J. H. 1997 *Influenial passengers. Inherited microorganisms and arthropod reproduction*. Oxford University Press.
- Sacchi, L., Grigolo, A., Laudana, U., Ricevuti, G. & Dealessi, F. 1985 Behavior of symbionts during oogenesis and early stages of development in the German cockroach, *Blattella germanica* (Blattodea). *J. Invert. Pathol.* **46**, 139–152.
- Sauer, C., Stackebrandt, E., Gadau, J., Hölldobler, B. & Gross, R. 2000 Systematic relationships and cospeciation of bacterial endosymbionts and their carpenter ant host species: proposal of the new taxon *Candidatus Blochmannia* gen. Nov. *Int. J. Syst. Evol. Microbiol.* **50**, 1877–1886.
- Schröder, D., Deppisch, H., Obermayer, M., Krohne, G., Stackebrandt, E., Hölldobler, B., Goebel, W. & Gross, R. 1996 Intracellular endosymbiotic bacteria of *Camponotus* species (carpenter ants): systematics, evolution, and ultrastructural characterization. *Mol. Microbiol.* **21**, 479–489.
- Swofford, D. L. 1998 *PAUP*: phylogenetic analysis using parsimony (and other methods)*, v. 4. Sunderland, MA: Sinauer Associates.
- Thao, L., Moran, N. A., Abbot, P., Brennan, E. B., Burckhardt, D. H. & Baumann, P. 2000 The speciation of psyllids and their primary prokaryotic endosymbionts. *Appl. Environ. Microbiol.* **66**, 2898–2905.