J Mol Evol (1996) 42:414–421

JOURNAL OF MOLECULAR EVOLUTION



# The Tryptophan Biosynthetic Pathway of Aphid Endosymbionts (*Buchnera*): Genetics and Evolution of Plasmid-Associated Anthranilate Synthase (*trpEG*) Within the Aphididae

**Dadbeh Rouhbakhsh,[1] Chi-Yung Lai,[1] Carol D. von Dohlen,[2] Marta A. Clark,[1] Linda Baumann,[1] Paul Baumann,[1] Nancy A. Moran,[2] David J. Voegtlin[3]**

[1] Microbiology Section, University of California, Davis, CA 95616-8665, USA
[2] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA
[3] Center for Biodiversity, Illinois Natural History Survey, Champaign, IL 61820-6970, USA

Received: 22 August 1995 / Accepted: 6 September 1995

**Abstract.** The bacterial endosymbionts (*Buchnera*) from the aphids *Rhopalosiphum padi, R. maidis, Schizaphis graminum,* and *Acyrthosiphon pisum* contain the genes for anthranilate synthase (*trpEG*) on plasmids made up of one or more 3.6-kb units. Anthranilate synthase is the first as well as the rate-limiting enzyme in the tryptophan biosynthetic pathway. The amplification of *trpEG* on plasmids may result in an increase of enzyme protein and overproduction of this essential amino acid, which is required by the aphid host. The nucleotide sequence of *trpEG* from endosymbionts of different species of aphids is highly conserved, as is an approximately 500-bp upstream DNA segment which has the characteristics of an origin of replication. Phylogenetic analyses were performed using *trpE* and *trpG* from the endosymbionts of these four aphids as well as from the endosymbiont of *Schlechtendalia chinensis,* in which *trpEG* occurs on the chromosome. The resulting phylogeny was congruent with trees derived from sequences of two chromosome-located bacterial genes (part of *trpB* and 16S ribosomal DNA). In turn, trees obtained from plasmid-borne and bacterial chromosome-borne sequences were congruent with the tree resulting from phylogenetic analysis of three aphid mitochondrial regions (portions of the small and large ribosomal DNA subunits, as well as cytochrome oxidase II). Congruence of trees based on genes from host mitochondria and from bacteria adds to previous support for exclusively vertical transmission of the endosymbionts within aphid lineages. Congruence with trees based on plasmid-borne genes supports the origin of the plasmid-borne *trpEG* from the chromosomal genes of the same lineage and the absence of subsequent plasmid exchange among endosymbionts of different species of aphids.

**Key words:** Aphid — *Buchnera aphidicola* — Cospeciation — Endosymbiosis — Mutualism — Tryptophan biosynthesis — *trpEG* — *trpB* — Gene amplification — Plasmid evolution

## Introduction

Symbiotic associations with intracellular prokaryotes are common in insects, especially within the sap-sucking species in the order Homoptera (Douglas 1989; Houk and Griffiths 1980). One well-studied endosymbiotic association is between *Buchnera* and aphids (Homoptera: Aphidoidea) (Baumann et al. 1995a,b; Douglas 1989; Ishikawa 1989). Most species of aphids contain, within their body cavity, a specialized bilobed structure (bacteriome) consisting of an aggregate of cells (bacteriocytes) which harbor *Buchnera.* The endosymbionts are housed in symbiosomes which are host-derived vesicles within bacteriocytes. A typical mature aphid (*Schizaphis graminum*) contains $5.6 \times 10^6$ cells of *Buchnera* (Baumann and Baumann 1994). The endosymbionts are transmitted ma-

*Correspondence to:* P. Baumann

ternally. The association between *Buchnera* and its host is mutualistic; neither can survive in the absence of the other.

Studies on the evolutionary relationships of *Buchnera* and comparisons with the phylogeny of the hosts support an infection of an aphid progenitor with a free-living bacterium 100–250 million years ago (Moran and Baumann 1994; Moran et al. 1993). Concordance between aphid and endosymbiont phylogenies supports subsequent cospeciation of endosymbionts and hosts, resulting in the present species of aphids and strains of *Buchnera.* These findings suggest exclusively vertical transmission of the endosymbionts with no exchange between aphids.

Insects are unable to synthesize a number of amino acids (Dadd 1985). Plant sap, the diet of aphids, is rich in carbohydrates but deficient in nitrogenous compounds, including certain essential amino acids (Douglas 1993). One of the functions postulated for *Buchnera* is the synthesis of these essential amino acids (Dadd 1985). The best evidence indicates that the endosymbionts are able to overproduce tryptophan, an amino acid that is not synthesized by the insects. Aphids are not able to grow on a synthetic diet lacking tryptophan if chlortetracycline is included (Douglas and Prosser 1992). Supplementation with tryptophan allows aphid growth. Since chlortetracycline eliminates endosymbionts, these results are interpreted as indicating that one of the functions of *Buchnera* is tryptophan biosynthesis. Consistent with this interpretation is the detection of tryptophan synthase activity (TrpBA, the last enzyme of the tryptophan biosynthetic pathway) in the endosymbionts and its elimination in aphids treated with chlortetracycline (Douglas and Prosser 1992).

*Schizaphis graminum* resembles other members of the Aphididae in having rapid development, reaching maturity 7–9 days after birth (Baumann and Baumann 1994). *Buchnera* from this aphid (*Buchnera*-Sg) contains all the genes of the tryptophan biosynthetic pathway (Lai et al. 1994; Munson and Baumann 1993). In *Buchnera*-Sg the genes for anthranilate synthase (TrpEG, the first enzyme of the pathway) are found on a plasmid which consists of four tandem repeats of a 3.6-kb *trpEG*-containing unit (Lai et al. 1994). The remaining genes of the pathway (*trpDC*[*F*]*BA*) are present as a single copy on the chromosome (Munson and Baumann 1993). Since there are about four plasmids per chromosome, the total *trpEG* amplification is 16-fold. Gene amplification is a common mechanism used by cells to increase the amount of enzyme protein under conditions when enzyme activity limits growth (Anderson and Roth 1977). The amplification of *trpEG* in *Buchnera*-Sg is an adaptation to a mutualistic association in which the endosymbiont overproduces tryptophan and supplies it to the aphid host.

In contrast to *S. graminum, Schlechtendalia chinensis* has a long developmental period, reaching maturity over a month after birth (Takada 1991). This relatively slow development is typical for the family Pemphigidae to which this species belongs. In the endosymbionts from this aphid (*Buchnera*-Sc), the genes of the tryptophan biosynthetic pathway are present as two linkage groups on the bacterial chromosome consisting of one copy each of *trpEG* and *trpDC(F)BA* (Lai et al. 1995). The lack of *trpEG* amplification in *Buchnera* (Sc) may be related to the relatively slow development of *S. chinensis,* which may impose a lower demand for tryptophan.

More recently, another example of plasmid-associated genes encoding for enzymes of amino acid biosynthesis has been found in aphid endosymbionts. Bracho et al. (1995) have shown that *Buchnera* from the aphid *Rhopalosiphum padi* (*Buchnera*-Rp) contains a plasmid with four genes of leucine biosynthesis. Evidence was also presented for the presence of this plasmid in *Buchnera* from several other species, all within the Aphididae.

In the present communication we examine the genetic organization of *trpEG* in *Buchnera* from *Rhopalosiphum padi*, *R. maidis* (Rm), and *Acyrthosiphon pisum* (Ap), three additional species of the family Aphididae which have rapid development (Dixon 1985). In order to test the hypothesis that the plasmid-borne *trp* genes are derived from the chromosome of the same lineage and subsequently are exclusively vertically inherited, we compare phylogenetic trees derived from *trpEG* with trees derived from genes on the bacterial chromosome (*trpB* and 16S ribosomal DNA [rDNA]) and a tree derived from genes on the host mitochondrial chromosome. If transmission is exclusively vertical, all trees should be congruent.

## Materials and Methods

Most of the methods used in this study have been described in detail in our past publications (Lai et al. 1994, 1995); only an outline will be presented here. The procedures used include purification of whole aphid DNA (containing both *Buchnera* and host genomes), restriction enzyme and Southern blot analysis, and nucleotide (nt) sequence determination. For all prokaryote sequences, both DNA strands were determined. Unless otherwise noted, standard methods in molecular biology were used (Sambrook et al. 1989).

*Cloning of the 3.6-kb* trpEG-*Containing Unit.* Using the synthetic oligonucleotides complementary to conserved regions of *trpE* previously described (Lai et al. 1995), in conjunction with the polymerase chain reaction (PCR), a 0.52-kb DNA fragment was amplified from each of the aphid-endosymbiont DNA preparations (*Buchnera*-Rp, -Rm, and -Ap) and ligated into pBluescript (Stratagene, La Jolla, CA). These plasmids were used as probes for the detection of *trpE* in restriction enzyme and Southern blot analyses of homologous DNA preparations. On the basis of these analyses restriction enzymes were identified for each DNA preparation which in Southern blots cut the *trpE*-containing DNA into 3.6-kb fragments. Enzymes compatible for cloning into *Eco*RI-cut, phosphatase-treated λZAP (Stratagene) were used to digest the DNA preparation (*Pvu*II for *R. padi*, *Eco*RV for *R. maidis*, *Eco*RI for *A. pisum*). The DNA was electrophoresed and the 3.4–3.8-kb fragments were electroeluted. *Eco*RI adaptors were added to the *Pvu*II- and *Eco*RV-digested DNA fragments and all preparations

were ligated into λZAP. The homologous 0.52-kb probe was used for the detection of recombinants. In all cases recombinants containing the 3.6-kb fragment in both orientations were cloned and sequenced; there were no differences between sequences of the two 3.6-kb units from *Buchnera* of the same aphid species.

*Cloning a Portion of* trpB. Using the synthetic oligonucleotides complementary to a portion of *trpB* described by Lai et al. (1995) and PCR, an approximately 0.7-kb amplified DNA fragment was cloned into pBluescript (Stratagene) in both orientations. The nt sequence was determined using a combination of T3 primers and the following synthetic oligonucleotide primers which were complementary to nt sequences approximately in the middle of the 0.7-kb *trpB* fragment: 5′GG(TA) CC(AG) CAT CCC TAT CC3′(downstream) and 5′CC(AT) CCA CC(AT) ACA CAT GC3′ (upstream).

*Amplification and Sequencing of Aphid Mitochondrial Regions.* PCR amplification was accomplished using synthetic oligonucleotide primers. Primers for COII amplified a 650-nt fragment from the 3′ end of COII to a position near the 5′ end of COII; primer sequences were 5′CAT TCA TAT TCA GAA TTA CC3′ (mt2993+ in Stern 1994) and 5′GAG ACC ATT ACT TGC TTT CAG TCA TCT3′ (similar to TK-N-3785 in Simon et al. 1994). In addition, a fragment spanning portions of s- and l-mitochondrial rDNA was amplified and sequenced using 5′AAA CTA GGA TTA GAT ACC CTA TTA T3′ (corresponding to SR-N-14588 in Simon et al. 1994) and 5′AGA TGA AAA CCA ACC TGG3′. Sequencing of aphid mitochondrial DNA followed methods of von Dohlen and Moran (1995). All of the COII fragment and portions of the rDNA fragment were sequenced with the exception that only part of COII was sequenced for *Melaphis rhois* (Mr) and *S. chinensis* and COII was not sequenced for the outgroup (*Adelges* sp.).

*Nt Sequence Accession Numbers.* The nt sequences were deposited in GenBank. The accession numbers are in parentheses following the aphid species designation: 3.6-kb plasmid unit: *R. padi* (L43551), *R. maidis* (L43550), and *A. pisum* (L43555); 0.7-kb portion of *trpB: R. padi* (L46358), *R. maidis* (L46356), *A. pisum* (L46355), and *M. rhois* (L46357); partial mitochondrial small subunit rDNA (s-rDNA): *R. padi* (U36736), *R. maidis* (U36735), *A. pisum* (U36733), *S. graminum* (U36738), *M. rhois* (U36734), *S. chinensis* (U36737), and *Adelges* sp. (U36732); partial mitochondrial large subunit rDNA (l-rDNA): *R. padi* (U36743), *R. maidis* (U36742), *A. pisum* (U36740), *S. graminum* (U36745), *M. rhois* (U36741), *S. chinensis* (U36744), and *Adelges* sp. (U36739); mitochondrial COII: *R. padi* (U36749), *R. maidis* (U36748), *A. pisum* (U36746), *S. graminum* (U36751), *M. rhois* (U36747), and *S. chinensis* (U36750). Sequences of mitochondrial rDNA were not obtained for the *Adelges* species.

*Phylogenetic Analysis.* Parsimony analyses were performed on all genes, using PAUP (Swofford 1993). Five analyses were performed: *trpE, trpG, trpB,* bacterial 16S rDNA, and a combined analysis for three aphid mitochondrial regions, almost the entire COII (650 nt) and sections from s-rDNA (344 nt) and l-rDNA (473 nt). For coding genes, amino acids were aligned first and DNA alignments made to fit; these alignments were unambiguous. For rDNA, alignments were performed using PILEUP of GCG (Genetics Computer Group 1994). Alignments of bacterial 16S rDNA were largely unambiguous. For the mitochondrial s-rDNA and mitochondrial l-rDNA, short regions in which the alignment was unclear were eliminated prior to the phylogenetic analysis. Additional *trp* nucleotide sequences used were *Buchnera*-Sg (Z19055, Z21938) and *Buchnera*-Sc (U09184, U09185). The outgroup for all of the bacterial gene analyses was *Vibrio parahaemolyticus* (*trp* genes X17149, 16S rDNA X74721), which is related to *Buchnera* and falls within the γ-3 subgroup of the *Proteobacteria.* The outgroup for the aphid mitochondrial analysis was a species of the genus *Adelges* (Homoptera: Adelgidae). Adelgidae, along with Phylloxeridae, is the most closely related group to aphids (von Dohlen and Moran 1995). For each of the 5 analyses, we ran 1,000 bootstrap replications using the heuristic search strategy option in PAUP. All sites were weighted equally. Total numbers of potentially informative sites for each analysis were 415 for *trpE,* 127 for *trpG,* 184 for *trpB,* 94 for 16S rDNA, and 86 for the combined aphid mitochondrial analysis.

*Substitution Rates at Synonymous and Nonsynonymous Sites in* trpB *and COII.* The divergence of *M. rhios* and *S. chinensis* can be approximately dated at 50–70 million years, based on biogeographic and paleobotanical evidence involving their specific host plants (Moran 1989). This estimate is consistent with molecular clock calculations using 16S rDNA sequences (Moran et al. 1993). We used this date to calibrate substitution rates at nonsynonymous sites and rates at synonymous sites for the partial *trpB* sequence and the partial COII region, using the method of Kumar et al (1994) in the program MEGA. Distances at each type of site were adjusted for multiple substitutions using the Jukes-Cantor correction.

*Function of the Putative* Buchnera *Plasmid Origin of Replication in* E. coli. *E. coli* JZ294 (*polA1*) is a mutant which is not able to support the replication of $ColE_1$ origins which are dependent on DNA polymerase I (Harding et al. 1982). $ColE_1$ replicons are contained in a variety of vectors including pBluescript (Stratagene). Competent cells of JZ294 were obtained by the $CaCl_2$ procedure (Sambrook et al. 1989) and transformed with pBluescript containing the 3.6-kb *trpEG* unit from *Buchnera*-Sg, -Rp, -Rm, and -Ap. In all of the 3.6-kb fragments the region between *trpG* and *trpE* was uninterrupted. As a control the cells were also transformed with (1) pIP637, which contains a pSC101 replicon (Huang et al. 1994) which is not dependent on DNA polymerase I, and (2) pBluescript. Transformants containing pIP637 were found, indicating that cells were competent; transformants containing pBluescript were not found, indicating that the strain of JZ294 had the correct phenotype.

pBSL15 is a vector containing a kanamycin resistance cassette (Alexeyev 1995). This gene was excised by digestion with *Eco*RI. In order to prevent religation to pBSL15, the preparation was also digested with *Hin*dIII. pBluescript containing the 3.6-kb unit from *Buchnera*-Sg and from *Buchnera*-Rm was digested with *Eco*RI, which resulted in excision of the 3.6-kb unit. Religation into pBluescript was prevented by digestion with *Xho*I. The two mixtures were combined and ligated; XLBlue (Stratagene) was transformed and plated on kanamycin-containing media (Alexeyev 1995).

## Results and Discussion

### *General Properties of the 3.6-kb Units*

A characteristic of circular plasmid DNA molecules, consisting of single units or identical tandemly repeated units, is that complete digestion of the DNA circle with restriction enzymes which cut once in the repeated units will give a DNA fragment of the same size. Restriction enzyme and Southern blot analysis of endosymbiont-aphid DNA preparations from *R. padi, R. maidis,* and *A. pisum,* using as a hybridization probe a homologous fragment of *trpE,* indicated that a variety of different enzymes digested the *trpE*-containing DNA to fragments of 3.6-kb (Baumann et al. 1995a,b). These 3.6-kb fragments were cloned and sequenced, and genetic maps, summarizing these results, are presented in Fig. 1. Each 3.6-kb unit contained two open reading frames (ORFs) corresponding to *trpE* and *trpG.* The nucleotide (nt) sequence

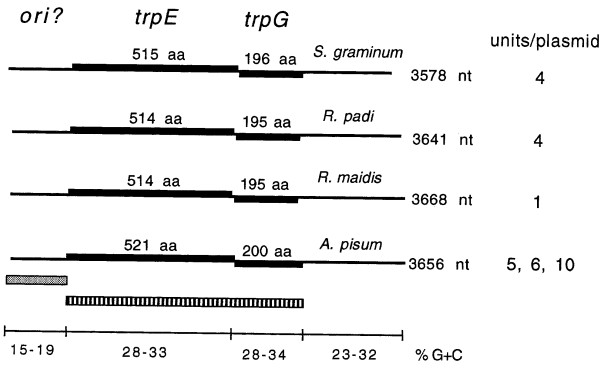


**Fig. 1.** Genetic maps of the 3.6-kb units which make up plasmids in some *Buchnera. Thin line,* untranslated DNA; *thick line,* genes coding for the two protein components of anthranilate synthase (*trpE, trpG*); *aa,* amino acid; *nt,* nucleotides; *stippled box,* conserved nt sequences of a putative origin of replication (*ori?*); *striped box,* conserved *trpE* and *trpG* nt sequences; *G + C,* guanine-plus-cytosine content of DNA segments. Data for *Buchnera* from *S. graminum* is from Lai et al. (1994).

corresponding to these genes was highly conserved; pairwise comparisons of the sequences indicated 70–88% nt identity. The nt sequence upstream of *trpE* (Fig. 1), which did not contain ORFs, was less conserved (50–76% nt identity). The nt sequence downstream of *trpG* showed no significant conservation (Fig. 1). The guanine-plus-cytosine (G + C) content of the 3.6-kb units ranged from 25–28 moles % for *Buchnera*-Sg, -Rp, and -Rm to 31 moles % for *Buchnera*-Ap. G + C content showed some regional differences, with the lowest found in the regions upstream of *trpE* (Fig. 1). The chromosomal *trpE* and *trpG* found in *Buchnera*-Sc (Lai et al. 1995) had 57–58% and 56–59% nt sequence identity, respectively, to homologous genes of *Buchnera*-Sg, -Rp, -Rm, and -Ap. In *Buchnera*-Sc, neither the region upstream of *trpE* nor the region downstream of *trpG* had significant nt sequence similarities to the corresponding regions of the plasmid-associated *trpEG.*

Another characteristic of circular plasmid DNA molecules, composed of identical tandem repeats, is that partial or timed digestions of the DNA will result in a ladder of fragments with sizes corresponding to multiples of the individual units (Lai et al. 1994). Timed digestions with *Xba*I, for *Buchnera*-Sg and *Pvu*II for *Buchnera*-Rp indicated the presence of fragments corresponding to 14.8, 10.8, 7.2, and 3.6 kb, corresponding to 4, 3, 2, and 1 units of 3.6-kb (Baumann et al. 1995a,b; Lai et al. 1994). *Buchnera*-Ap contained a mixture of plasmid sizes corresponding to 5, 6, and 10 units (Baumann et al. 1995a,b). It is not known if this variation in number of units per plasmid occurs within individual endosymbionts or within individual aphids. *Buchnera*-Rm differed from the other endosymbionts in having a plasmid consisting of one unit (Baumann et al. 1995a,b). These results are compiled in Fig. 1 and are interpreted as indicating that amplification of *trpEG* can occur by an increase in the number of tandem repeats which are present in a plasmid as well as by an increase in the plasmid copy number (Baumann et al. 1995a,b).

### *Sequence Conservation of TrpE and TrpG*

Figure 2 presents the deduced, aligned amino acid sequences corresponding to TrpE and TrpG of *Buchnera*-Rp, -Rm, and -Ap as well as the previously published sequences of these two proteins from *Buchnera*-Sg and *Buchnera*-Sc (Lai et al. 1994, 1995). The number of residues identical for all taxa was 271 for TrpE and 103 for TrpG. Comparisons of amino acid sequences among the different endosymbionts indicated identities ranging from 55 to 89% for TrpE and from 59 to 93% for TrpG.

### *Conserved nt Sequences Upstream of* trpE

The sequences of the aligned 508–600 bp nt regions upstream of *Buchnera trpE* (Fig. 1) are presented in Fig. 3. The nt identity of the sequences of *Buchnera*-Sg, -Rp, and -Rm was 69–75%. The *Buchnera*-Ap sequence had 50–55% nt identity to the corresponding region of these three endosymbionts. These high-sequence similarities, together with the absence of significant ORFs, suggest that this region has important functional attributes. In order to maintain itself as a plasmid, a circular piece of DNA must have an origin of replication (*ori*) (Maloy et al. 1994). In many cases, the characteristics of an *ori* are one or more copies of a nine-nt sequence known as a *DnaA box* (consensus, TTGTCCACA) as well as a region containing a low G + C content (Yoshikawa and Ogasawara 1991). The *DnaA box* is the region to which DnaA protein binds while the region of low G + C content is the site at which the two strands come apart and replication begins (Georgopoulos 1989). A portion of the nt sequence presented in Fig. 3 has these properties in that one consensus *DnaA box* is followed by two potential *DnaA boxes,* in an opposite orientation, which differ in 1 nt from the consensus sequence (2 nt in the case of one *Buchnera*-Ap *DnaA box*) (Fig. 3, lines 100 and 200). *Buchnera*-Sg contains two 25-nt direct repeats (Fig. 3, line 400). The sequence corresponding to only one of these direct repeats is conserved in all four endosymbionts, suggesting that the direct repeat observed in *Buchnera*-Sg is due to a duplication. This sequence contains a stretch of alternating A's and T's and has a low G + C content. The conserved sequences beginning at nt 511 could correspond to the *trp* promoter. The sequence TTGATT—14 nt—AGTTAT has some similarity to the putative −35 and −10 regions preceding *Buchnera* ribosomal RNA genes (Munson et al. 1993; Rouhbakhsh and Baumann 1995). The sequence GAG (nt 600–602), 10 nt upstream of the ATG which initiates *trpE,* may be part of a ribosome binding site.

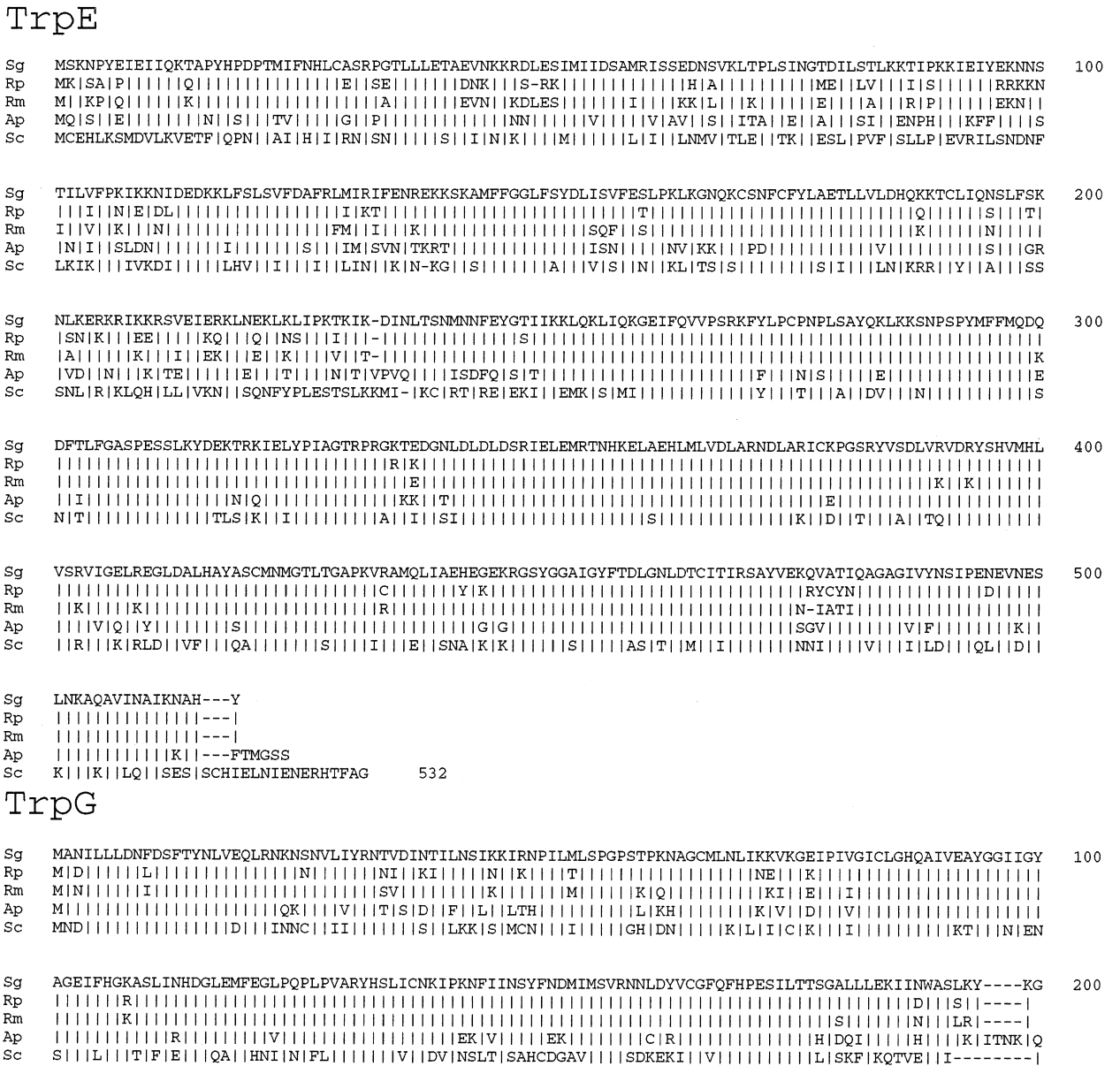


**Fig. 2.** Amino acid sequences of TrpE and TrpG from *Buchnera* of five species of aphids. |, amino acid identical to the one above; −, absence of an amino acid. Aphid hosts: *Sg, S. graminum; Rp, R. padi; Rm, R. maidis; Ap, A. pisum; Sc, S. chinensis.* Data for Sg and Sc from Lai et al. (1994, 1995), respectively.

### *Phylogenetic Relationships*

Each of the five maximum parsimony analyses resulted in a single shortest tree (Fig. 4). These trees are essentially identical, irrespective of the molecule chosen for analysis. The sole difference is in the order of branching of *Buchnera*-Sg, -Rp, and -Rm. Relationships among these three species could not be adequately resolved.

Congruence between bacterial trees and the aphid mitochondrial tree adds to previous evidence suggesting exclusively vertical transmission of *Buchnera* within aphid lineages (Moran and Baumann 1994). In *Buchnera*-Sg, -Rp, -Rm, and -Ap, *trpEG* is plasmid-borne, whereas in *Buchnera*-Sc the genes for *trpE* and *trpG* are in the ancestral chromosomal position (Lai et al. 1995). Congruence between trees based on *trpEG* and trees based on genes on the bacterial chromosome and aphid mitochondrial chromosome indicates that plasmid-associated *trpEG* is derived from the ancestral chromosomal genes of the same lineage rather than from an exogenous source. This congruence also supports exclusively vertical transmission of the plasmic-borne genes and the absence of transfer of plasmids among *Buchnera* lineages. This finding contrasts with the situation in free-living bacteria, in which plasmids are often readily exchanged between unrelated bacterial hosts, resulting in

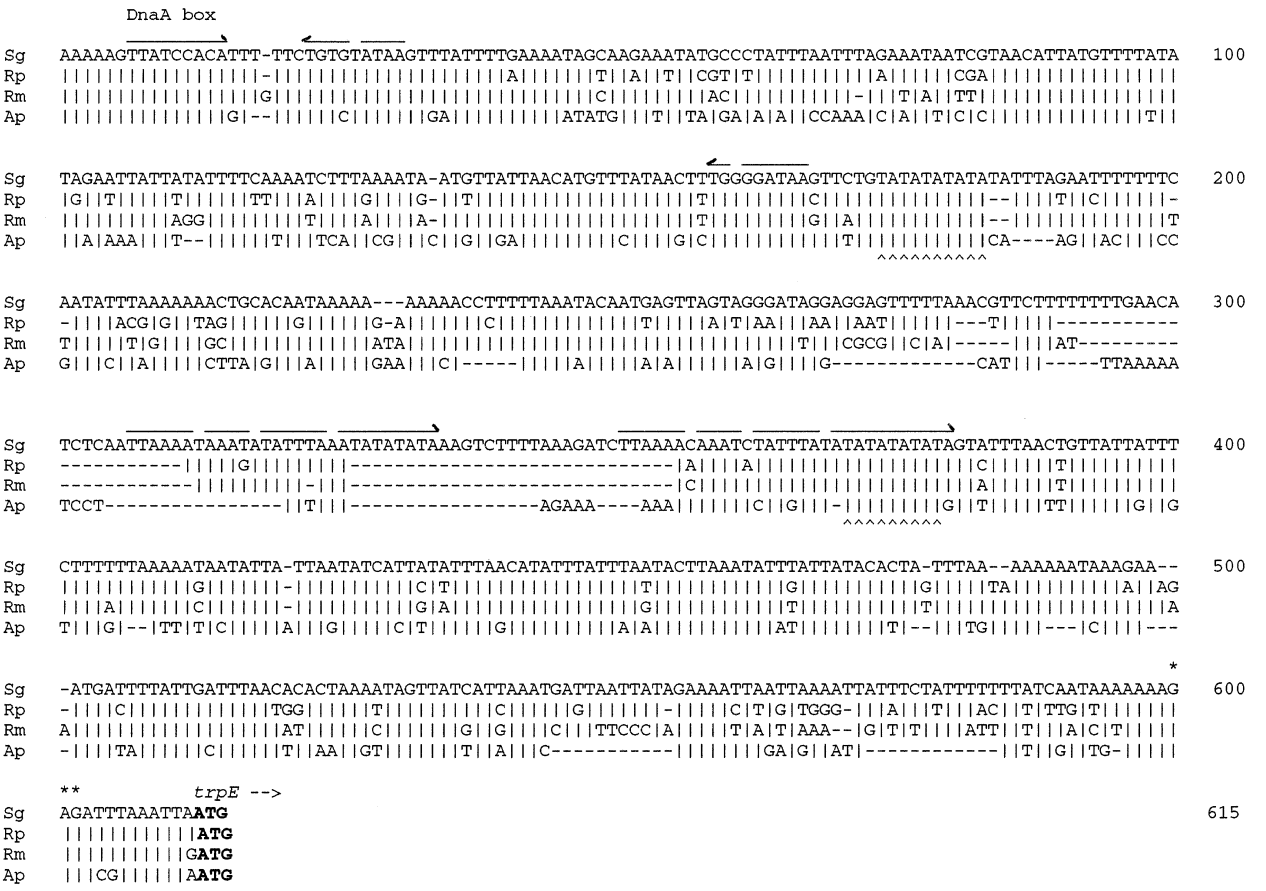


**Fig. 3.** Conserved nt sequences corresponding to the putative origin of replication of plasmids from *Buchnera* from four species of aphids. *Arrows above* the nt sequences, inverted repeats and direct repeats (the 25-nt direct repeat is only present in Sg); |, nt identical to the one above; ^, row of alternating TAs; *, portion of a putative ribosome binding site (nt 600–602).

horizontal transfer of plasmid-borne attributes (Campbell 1981).

The results presented in Fig. 4 are also consistent with the classification of the aphid hosts (Blackman and Eastop 1984). *S. graminum, R. padi,* and *R. maidis* are in the tribe Aphidini; *A. pisum* is in the tribe Macrosiphini. Both of these tribes are in the family Aphididae. *M. rhois* and *S. chinensis* are both in the tribe Fordini within the family Pemphigidae.

### *Rates of Substitution in* trpB *and COII*

Synonymous substitutions can be considered to approximate neutrality with respect to fitness and so might be expected to evolve in a clock-like manner over time. The corrected number of substitutions per site for synonymous sites for *M. rhois* − *S. chinensis* is 0.678 ± 0.152 substitution/site for partial COII (mitochondrial) and 0.900 ± 0.131 for the partial *trpB* (bacterial). Calibrating with 60 MY as the estimated date of divergence for these two species (Moran et al. 1993), these distances correspond to rates of approximately 0.44–0.69 substitutions/site/100 MY for COII and 0.64–0.86 substitutions/site/100 MY for *trpB.* Thus, there appears to be no substantial difference in rates of neutral evolution between host mitochondrial genes and bacterial genes. In contrast, rates of substitution in host nuclear 18S rDNA are much slower than rates in homologous parts of the bacterial 16S rDNA (Moran et al. 1995).

### *Function of* Buchnera ori *in* E. coli

pBluescript vectors contain a $ColE_1$ origin of replication which is unique among bacterial origins of replication in that it is dependent on DNA polymerase I (Kingsbury and Helinski 1973). An *E. coli* mutant defective in this enzyme is not able to support replication of pBluescript containing the 3.6-kb *trpEG*-containing inserts, indicating that the putative *Buchnera ori* is not able to function in this organism. Attempts to detect replication in *E. coli* XLBlue (Stratagene) of the 3.6-kb fragments ligated to a gene for kanamycin resistance from pBSL15 (Alexeyev 1995) were unsuccessful. These results indicate that the putative *Buchnera ori* region does not function in this capacity in *E. coli.*

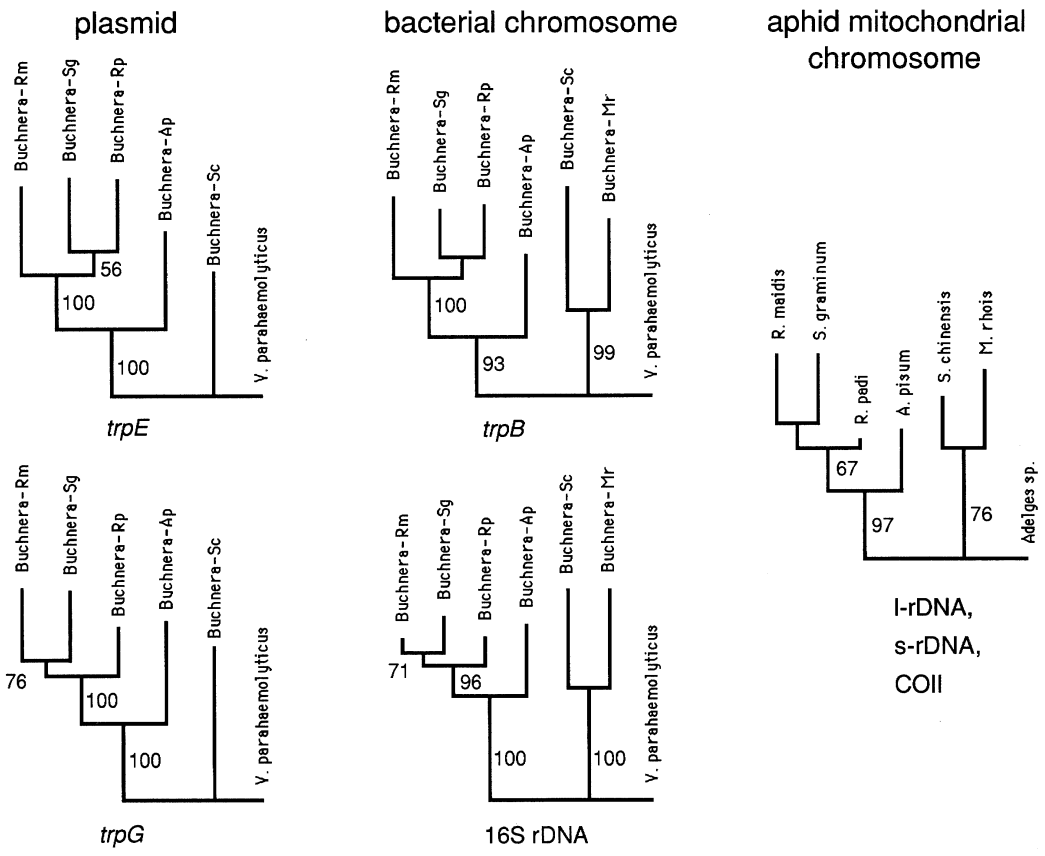


**Fig. 4.** Phylogenetic trees resulting from parsimony analyses of nt sequences of *trpE, trpG, trpB* (partial), 16S rDNA from *Buchnera* and three regions of the aphid mitochondrial chromosome. *Abbreviations* indicating the aphid hosts are identified in the text. *Numbers* are bootstrap percentages from parsimony searches (1,000 replicates), presented only for nodes with greater than 50% support.

## Conclusions

From present and past investigations, the following conclusions can be drawn:

1. In four representatives of Aphididae (*S. graminum, R. padi, R. maidis,* and *A. pisum*), *trpEG* is plasmid-associated. In *S. chinensis,* a member of Pemphigidae, *trpEG* is present on the chromosome. Species in Aphididae show generally faster development than most Pemphigidae, including *S. chinensis.* The plasmid location and amplification of *trpEG* may be connected with a higher demand for tryptophan by species in Aphididae.
2. Phylogenetic analysis of *Buchnera* plasmid-borne and chromosomal *trpE* and *trpG,* of chromosomal *trpB* and 16S rDNA, and of aphid mitochondrial genes indicates congruence between the trees derived from these molecules. These results confirm the long-term vertical transmission of *Buchnera,* in agreement with previous phylogenetic evidence (Moran et al. 1993). Agreement between trees based on plasmid-borne genes and trees based on chromosomal and mitochondrial genes indicates that the plasmid-borne *trpEG* is derived from the ancestral chromosomal copy and not an exogenous source and that endosymbionts of different aphid species do not exchange plasmids.
3. A DNA segment found in each of the 3.6-kb *trpEG*-containing units has nt sequences characteristic of an origin of replication. DNA fragments containing this segment do not function as an origin of replication in *E. coli.*

*Acknowledgments.* *E. coli* JZ294 and plasmids pIP637 and pBSL15 were the gifts of J. Zyskind, M.M. Igo, and M.F. Alexeyev, respectively. This material is based on work supported by the National Science Foundation under awards No. MCB-9402813 and DEB-9306495 and by the University of California Agricultural Experiment Station.

## References

Alexeyev MF (1995) Three kanamycin resistance gene cassettes with different polylinkers. Biotechniques 18:52–54

Anderson RP, Roth JR (1977) Tandem genetic duplications in phage and bacteria. Annu Rev Microbiol 31:473–505

Baumann P, Baumann L (1994) Growth kinetics of the endosymbiont, *Buchnera aphidicola,* in the aphid *Schizaphis graminum.* Appl Env Microbiol 60:3440–3443

Baumann P, Baumann L, Lai C-Y, Rouhbakhsh D, Moran NA, Clark

MA (1995a). Genetics, physiology, and evolutionary relationships of the genus *Buchnera:* intracellular symbionts of aphids. Annu Rev Microbiol 49:55–94

Baumann P, Lai C-Y, Baumann L, Rouhbakhsh D, Moran NA, Clark MA (1995b) Mutualistic association of aphids and prokaryotes: biology of the genus *Buchnera.* Appl Environ Microbiol 61:1–7

Blackman RL, Eastop VF (1984) Aphids on the World's Crops. Wiley, Chichester

Bracho AM, Martínez-Torres D, Moya A, Latorre A (1995) Discovery and molecular characterization of a plasmid localized in *Buchnera* sp., bacterial endosymbiont of the aphid *Rhopalosiphum padi.* J Mol Evol 41:67–73

Campbell A (1981) Evolutionary significance of accessory DNA elements in bacteria. Annu Rev Microbiol 35:55–83

Dadd RH (1985) Nutrition: organisms. In: Kerkut GA, Gilbert GI (eds) Comprehensive insect physiology biochemistry and pharmacology, vol. 4. Pergamon, Elmsford, pp 315–319

Dixon AFG (1985) Aphid ecology. Blackie, London

Douglas AE (1989) Mycetocyte symbiosis in insects. Biol Rev Camb Philos Soc 64:409–434

Douglas AE (1993) The nutritional quality of phloem sap utilized by natural aphid populations. Ecol Entomol 18:31–38

Douglas AE, Prosser WA (1992) Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrthosiphon pisum*) symbiosis. J Insect Physiol 38:565–568

Genetics Computer Group (1994) Program manual for the Wisconsin package, version 8. University of Wisconsin, Madison, Wisconsin

Georgopoulos C (1989) The *E. coli dnaA* initiation protein: a protein for all seasons. Trends Genet 5:319–321

Harding NE, Cleary JM, Smith DW, Michon JJ, Brusilow WSA, Zyskind JW (1982) Chromosomal replication origins (*oriC*) of *Enterobacter aerogenes* and *Klebsiella pneumoniae* are functional in *Escherichia coli.* J Bacteriol 52:983–993

Houk EJ, Griffiths GW (1980) Intracellular symbiotes of the Homoptera. Annu Rev Entomol 25:161–187

Huang K-J, Schieberl JL, Igo MM (1994) A distant upstream site involved in the negative regulation of the *Escherichia coli ompF* gene. J Bacteriol 176:1309–1315

Ishikawa H (1989) Biochemical and molecular aspects of endosymbiosis in insects. Int Rev Cytol 116:1–45

Kingsbury DT, Helinski DR (1973) Temperature-sensitive mutants for the replication of plasmids in *Escherichia coli:* requirement for deoxyribonucleic acid polymerase I in the replication of plasmid $ColE_1$. J Bacteriol 114:1116–1124

Kumar S, Tamura K, Nei M (1994) MEGA, Molecular evolutionary genetics analysis, version 1.01. Institute of Molecular Evolutionary Genetics. The Pennsylvania State University, University Park, PA

Lai CY, Baumann L, Baumann P (1994) Amplification of *trpEG;* adaptation of Buchnera aphidicola to an endosymbiotic association with aphids. Proc Natl Acad Sci USA 91:3819–3823

Lai CY, Baumann P, Moran NA (1995) Genetics of the tryptophan biosynthetic pathway of the prokaryotic endosymbiont (*Buchnera*) of the aphid *Schlechtendalia chinensis.* Insect Mol Biol 4:47–59

Maloy SR, Cronan JE, Freifelder D (1994) Microbial genetics. Jones and Bartlet, Boston

Moran NA (1989) A 48-million-year-old aphid-host plant association and complex life cycle:biogeographic evidence. Science 245:173–175

Moran NA, Baumann P (1994) Phylogenetics of cytoplasmically inherited microorganisms of arthropods. Trends Ecol Evol 9:15–20

Moran NA, Munson MA, Baumann P, Ishikawa H (1993) A molecular clock in endosymbiotic bacteria calibrated using the insect hosts. Proc R Soc Lond [Biol] 253:167–171

Moran NA, von Dohlen CD, Baumann P (in press) Faster evolutionary rates in endosymbiotic bacteria than in cospeciating insect hosts. J Mol Evol 41:727–731

Munson MA, Baumann L, Baumann P (1993) *Buchnera aphidicola* (a prokaryotic endosymbiont of aphids) contains a putative 16S rRNA operon unlinked to the 23S rRNA-encoding gene: sequence determination, and promoter and terminator analysis. Gene 137:171–178

Munson MA, Baumann P (1993) Molecular cloning and nucleotide sequence of a putative *trpDC(F)BA* operon in *Buchnera aphidicola* (endosymbiont of the aphid *Schizaphis graminum*). J Bacteriol 175: 6426–6432

Rouhbakhsh D, Baumann P (1995) Characterization of a putative 23S–5S rRNA operon of *Buchnera aphidicola* (endosymbiont of aphids) unlinked to the 16S rRNA-encoding gene. Gene 155:107–112

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning; a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (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

Stern DL (1994) A phylogenetic analysis of soldier evolution in the aphid family Hormaphidinae. Proc R Soc Lond [Biol] 256:203–209

Swofford DL (1993) PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, IL

Takada H (1991) Gall development of *Schlechtendalia chinensis* (Bell) (Homoptera: Pemphigidae) on Rhus javanica L. and emergence of alatae from galls. Jpn J Appl Entomol Zool 35:71–76

von Dohlen CD, Moran NA (1995) Molecular phylogeny of the Homoptera: a paraphyletic taxon. J Mol Evol 41:211–223

Yoshikawa H, Ogasawara N (1991) Structure and function of DnaA and the DnaA-box in eubacteria: evolutionary relationships of bacterial replication origins. Mol Microbiol 5:2589–2597