The Endosymbiont (Buchnera sp.) of the Aphid Diuraphis noxia Contains Plasmids Consisting of trpEG and Tandem Repeats of trpEG Pseudogenes

CHI-YUNGLAI,1 PAUL BAUMANN,1* AND NANCY MORAN2

Microbiology Section, University of California, Davis, California 95616-8665,1 and Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 857212

Received 18 September 1995/Accepted 7 November 1995

Most aphids are dependent for their survival on prokaryotic endosymbionts assigned to the genus Buchnera. Among the functions of Buchnera species is the synthesis of tryptophan, which is required by the aphid host. In Buchnera species from the aphid Diuraphis noxia, the genes for anthranilate synthase (trpEG) were found on a plasmid which consisted of seven tandem repeats of a 3.2-kb unit and one 2.6-kb unit which differed in containing a 0.6-kb deletion. One of the 3.2-kb units contained open reading frames corresponding to trpEG; the remaining units contained trpEG pseudogenes (ϕ). The nucleotide sequence upstream of trpE contained a region that has characteristics of an origin of replication (ori). Relative to trpB (a chromosomal gene), there were about two copies of the trpE-containing plasmid. Comparisons of the nucleotide sequences of the 3.2-kb units containing trpEG and ϕtrpEG indicated that most changes occurred in a 700-nucleotide segment that included the region upstream of trpE and the portion of this gene coding for the N terminus. The consequence of these changes was the silencing of trpEG by inactivation of the putative promoter region and premature termination of the TrpE peptide. In contrast, the nucleotide sequence of the segment corresponding to ori was conserved in the units containing trpEG and ϕtrpEG. We offer a number of speculations on the evolutionary pressure in this lineage which resulted in the silencing of most of trpEG while still retaining the regions resembling ori.

Aphids are major agricultural pests, causing substantial economic losses (7, 37). They feed by penetrating the plant by means of a tube (styler) which reaches the phloem tissue and serves as a conduit for plant sap (12). Most species of aphids contain endosymbionts (Buchnera species), which are present within specialized cells (bacteriocytes) (4, 5, 13, 18, 19). These cells form a loose bilobed aggregate, called the bacteriome, which is found within the aphid body cavity. The endosymbionts and the aphid are dependent on each other. Elimination of Buchnera cells by various treatments lead to sterility and eventual death of the aphid; the endosymbionts have not been cultured outside the aphid host. During their most active reproductive state, aphids are females which reproduce parthenogenetically, giving birth to live young (12). When mature, Schizaphis graminum, a typical aphid, weighs about 0.5 mg and contains 5.6 × 10^6 Buchnera cells (3).

Phloem sap is a diet rich in carbohydrates but poor in nitrogenous compounds such as amino acids, which are required by insects (11). Perhaps the best-documented function of the endosymbiont is the synthesis of the essential amino acid tryptophan, which is provided for the aphid host (4, 5, 14, 21). Nutritional experiments have shown that aphids cannot grow on a complex synthetic diet containing chlortetracycline and lacking tryptophan (14). The addition of tryptophan restores a limited growth capability. These experiments suggest that elimination of the endosymbiont by the antibiotic results in a requirement for exogenously supplied tryptophan. In addition, tryptophan synthase, the last enzyme of the tryptophan biosynthetic pathway, has been detected in endosymbionts but is absent from aphids treated with chlortetracycline (14).

Anthranilate synthase (TrpEG) is the first enzyme of the tryptophan biosynthetic pathway, and its activity controls the rate of tryptophan biosynthesis (10, 21). Recently, Buchnera species from the aphids Schizaphis graminum, Rhopalosiphum padi, R. maidis, and Acrithosiphon pisum were found to have trpEG on plasmids (4, 5, 21, 32). All of these aphids have a short development time and are in the family Aphididae. The trpEG-containing plasmids consisted of 3.6-kb units which were present as a single copy or as 4 to 10 tandem repeats. Amplification of trpEG would provide a means of increasing the amount of enzyme protein and result in the overproduction of tryptophan for the aphid host (21). In free-living bacteria, gene amplification by tandem duplication is a common mechanism for increasing the level of an enzyme whose activity limits growth (1, 6). When the selective pressure is removed, there is usually a rapid decrease of the gene copy number (1, 6). Unlike these four species of the Aphididae, the aphid Schlechtendalia chinensis, in the family Pemphigidae, has a long development time (23). Its demand for tryptophan is presumably low, and Buchnera species from this aphid have trpEG as a single copy on the endosymbiont chromosome (23). In Buchnera species from Schizaphis graminum and Schlechtendalia chinensis, the remaining genes of the tryptophan biosynthetic pathway [trpDC(F)BA] are located on the endosymbiont chromosome and probably constitute a single transcription unit (21, 23, 29).

Diuraphis noxia (Russian wheat aphid) is an aphid with a short development time and is a pest of grains. This aphid was previously restricted to southern Russia, countries bordering the Mediterranean, Iran, and Afghanistan (20, 31). In 1978, it was found as a pest of wheat in the Republic of South Africa. In 1986, it was detected in Texas and subsequently it rapidly spread to 14 other states and Canada, causing major damage to wheat and barley (20, 31). D. noxia is in the family Aphididae.
Evolutionary studies with 16S rDNA have shown that the Buchnera species that is associated with D. noxia is closely related to the Buchnera species associated with Schizaphis graminum, R. padi, R. maidis, and A. pisum, with the last being its closest relative among these species (27). These results are consistent with the classification of the aphid hosts (7).

The damage caused by D. noxia on wheat is markedly different from and much more drastic than that caused by R. padi and Schizaphis graminum, two other pests of wheat (9, 16). The plant damage characteristic of D. noxia includes leaf rolling, longitudinal white leaf streaks, purple discoloration, and prostrate growth. Histological studies indicate that the first signs of damage involve chloroplast membranes. This is followed by disintegration of chloroplasts as well as of other organelles, with the resulting cell being empty and devoid of visible structure (16). Compared with the damage caused to the plant by D. noxia, that caused by R. padi and Schizaphis graminum is milder, since these aphids do not cause complete disintegration of cell organelles (16). As with other plants, the phloem of wheat has a low concentration of tryptophan (15). Since feeding by D. noxia causes major disintegration of plant tissue, we reasoned that perhaps free tryptophan is available for the aphid. The resulting reduction in the demand for tryptophan biosynthesis may have affected trpEG amplification by modifying the number of tandem repeats and/or the plasmid copy number. In this study, we have examined the trpEG-containing plasmid of the Buchnera species associated with D. noxia. Our results indicate that the principal endosymbiotic plasmid contains one trpEG unit and, unexpectedly, approximately seven tandem repeats of a trpEG pseudogene.

MATERIALS AND METHODS

General methods. Standard molecular biology methods were used (2, 34). Details of these procedures, as well as slight modifications, have been described in our past publications (21, 29). The methods include the purification of endosymbiont-DNA and restriction enzyme and Southern blot analysis (Genius system; Boehringer Mannheim Corp., Indianapolis, Ind.), and sequencing of DNA by the dideoxy-chain termination method with 35S-dATP. In all cases, both DNA strands were sequenced.

Cloning of the trpEG-containing fragments. Similar procedures were previously described for a variety of Buchnera genes (4, 29). Only a brief outline of these methods is given here. Using the oligonucleotide primers to trpE described previously (23), a 0.52-kb DNA fragment was amplified from the endosymbiont-D. noxia DNA preparation by PCR and cloned into pBluescript (Stratagene, La Jolla, Calif.) (32). About half of the insert was sequenced to confirm its identity as trpE. This 0.52-kb fragment was used as a probe in the initial restriction enzyme and Southern blot analysis. On the basis of these results, as well as subsequent nucleotide sequence determinations, we decided to clone into ZAP (Stratagene) the following DNA fragments (see Fig. 1 and 2): (i) 1.6-kb EcoRI-EcoRV, (ii) 2.6-kb HindIII, (iii) 2.6-kb HpaI, and (iv) 3.2-kb HindIII. The D. noxia-Buchnera DNA preparation was digested with the indicated enzymes and subjected to agarose gel electrophoresis. The fragments in the expected size range were cut out of the agarose gel, electroeluted, and ligated to the appropriate adaptors or linkers containing an EcoRI site. After digestion with EcoRI, the DNA was ligated to phosphate-treated ZAP and bacteriophage containing the recombimants was detected by plaque hybridization with the 0.52-kb probe to trpE. The DNA was excised into pBluescript and sequenced.

After determination of the sequences of some of the inserts, a 0.33-kb probe to trpE was used in one subsequent cloning experiment and in restriction enzyme and Southern blot analysis (see Fig. 2 to 4). This probe was obtained by PCR with plasmids with the 2.6-kb HindIII trpEG-containing fragment and the following oligonucleotide primers: 5'-GTT ATC TTC AGC GCC G-3' and 5'-CTG AAA ACC ACA CAC TCG-3'. Since HindIII and EcoRV sites were found in this 0.33-kb probe, it was useful for the detection of DNA fragments adjacent to both sides of these restriction sites.

We were not able to detect recombinants containing the 1.7-kb HindIII-EcoRI fragment (see Fig. 2) by using the ZAP-pBluescript vector system. Previous work has shown that DNA containing trpE and a putative trpG gene on plasmids of replication (ori) was toxic to Escherichia coli when present in pBluescript, a high-copy-number vector (21). Therefore, our inability to clone the 1.7-kb HindIII-EcoRI fragment could be due to its toxicity in a multicopy plasmid. On the basis of our deductions from the nucleotide sequence and the restriction enzyme analysis (see Fig. 1 and 2), we made the following oligonucleotide primers, which were complementary to sequences upstream of HindIII and downstream of EcoRI: 5'-CTG AAA ATG AGG GAA TTC-3' and 5'-GTT CTC TCC AAG GCC G-3'. The DNA was amplified by PCR, digested with HindIII and EcoRI, and ligated into the low-copy-number vectors pWSK129 and pWSK130 (38) (gift of S. R. Kushner), and the nucleotide sequence was determined.

Time course of digestion of DNA containing trpEG. The reaction mixture at each time point contained 3 μg of D. noxia-Buchnera DNA with 2 U of HindIII or EcoRV per μg of DNA. Electrophoresis was performed on 0.6% agarose gels and was followed by Southern blotting. The 0.33-kb probe to trpE was used. The molecular weight standards used in the Southern blots were HindIII- and HinfIII-EcoRI-digested λ bacteriophage. These standards, as well as 13 DNA fragments spanning the range of 8.2 to 48.5 kb, were also used in a portion of the agarose gel that was stained with ethidium bromide. The results are expressed as averages of determinations obtained from four replicate Southern blots.

Evolutionary analyses. Gene phylogenies were reconstructed on the basis of nucleotide sequences of trpE and trpG and of the corresponding pseudogenes from the 3.2- and 2.6-kb units. Also included in the analyses were trpEG from Buchnera species associated with A. pisum and with Schizaphis graminum. Schizaphis graminum served as outgroup, based on previous data about species relationships (27). We reconstructed gene phylogenies by using parsimony (36) with all nucleotide sites included and weighted equally. The pseudogene sequences included insertions and deletions; these sites in the alignment were counted as missing in the phylogenetic analysis.

Nucleotide sequence accession number. The nucleotide sequence of the 0.034-bp DNA fragment has been deposited in GenBank under accession no. L46769.

RESULTS

Overview. To facilitate presentation, we will give a summary of the sequence data followed by the results which led to our interpretation. Figure 1 shows a genetic and restriction site map of the 9,034-bp fragment assembled from the nucleotide sequence of the cloned fragments. The G+C content of this DNA fragment is 29.0 mol%, a value consistent with the Buchnera G+C content (28 mol%) (4). The first two significant open reading frames correspond to trpE and trpG. The nucleotide sequences of D. noxia-associated Buchnera trpE and trpG are both 75% identical to the corresponding A. pisum-associated Buchnera genes (this species is the closest relative of the D. noxia-associated Buchnera species) (29, 32). The deduced amino acid sequences of TrpE and TrpG are 74 and 77% identical, respectively, to those of these proteins from Buchnera species associated with A. pisum. Downstream from trpEG are nucleotide sequences containing segments with a high similarity to trpE and trpG. These DNA segments do not constitute open reading frames, since they contain numerous nucleotide changes preventing the synthesis of a functional protein, and they are therefore inactive genes, or pseudogenes, identified by the letter ψ preceding the designation of the gene which they resemble (Fig. 1). The results of the nucleotide sequence and restriction enzyme analysis of endosymbiont DNA (Fig. 1) have been interpreted to indicate that Buchnera trpEG is present on plasmids consisting of (i) 3.2-trp, 3.2-kb tandem repeats of a ψtrpEG-containing fragment which lacks EcoRI sites; (ii) 3.2trp, a 3.2-kb trpEG-containing fragment with an EcoRI site; and (iii) 2.6-trp, an adjacent 2.6-kb ψtrpEG-con-
taining fragment lacking EcoRI sites, which differs from 3.2Δtrp in having a 555-bp deletion. The EcoRI site present in 3.2Δtrp is in the trpE coding region. Its absence in 3.2Δtrp and 2.6Δtrp is due to the change of two nucleotides, from GAATTC to AAATTT.

**Restriction enzyme analysis.** A characteristic of DNA circles consisting of tandem repeats of the same unit is that digestion with different restriction enzymes which cut once in the repeated unit will give the same fragment size (21). This was the case with trpEG-containing plasmids from Buchnera species from four species of aphids (21, 32). Restriction enzyme and Southern blot analysis led to a similar but slightly different conclusion for Buchnera species from *D. noxia* (Fig. 2). Single digestions with HindIII, EcoRV, Clal, SalI, and HpaI result in an intense band of 3.2 kb and a lighter band of 2.6 kb (Fig. 2, lanes 1, 3, 6, 8, 10), a finding consistent with multiple tandem repeats of a 3.2-kb unit and fewer (or one) related fragments of 2.6 kb (for an example, see Fig. 1, HindIII fragments). Digestion with each of these enzymes in combination with EcoRI results in two additional fragments which add up to 3.2 kb. For the combination of EcoRI and HindIII or EcoRV, because of the position of the probe (Fig. 1), both of these fragments are detected (Fig. 2, lanes 2 and 4). The remaining detected fragments of combined digestion with EcoRI and either Clal, SalI, or HpaI are consistent with the position of the probe and the position of the restriction site in the adjacent (left) 3.2Δtrp fragment (Fig. 1 and 2, lanes 7, 9, and 11).

**Estimation of the number of tandem repeats.** Southern blot analysis of undigested *D. noxia*-associated *Buchnera* DNA showed the presence of one major band and two minor bands migrating at a position above the 48.5-kb molecular size standard (Fig. 3A, lane 1); in addition, there were minor bands corresponding to approximately 27, 24, and 21 kb. Digestion with EcoRI resulted in elimination of the three bands above 48.5 kb and the formation of a major band at 24 kb and two minor bands at 27 and 21 kb (lane 2). By analogy with past experiments (21, 32), these results suggest that the bands above 48.5 kb correspond to relaxed DNA circles, which are linearized upon digestion of a single EcoRI site. The sum of the results obtained from restriction enzyme analysis and the nucleotide sequence (Fig. 1 and 2) is consistent with *D. noxia*-associated *Buchnera* plasmids having one 3.2trp-2.6Δtrp segment; the major plasmid also contains six tandem repeats of 3.2Δtrp, and two minor plasmids contain five and seven tandem repeats of 3.2Δtrp. Such plasmids consisting of one 2.6-kb unit and six, seven, or eight 3.2-kb units would have sizes corresponding to 21.8, 25.0, and 28.2 kb, respectively. These calculated values are slightly larger than the sizes of the EcoRI-digested, linearized plasmids estimated from Southern blots (Fig. 3A, lane 2). Timed digestions of the DNA preparation with restriction enzymes were also performed to further elucidate the arrangement of the tandem repeats. Upon digestion with HindIII, there are no intermediates between the high-molecular-mass bands and the 27-, 24-, and 21-kb bands (Fig. 3B, lanes 3 and 4), consistent with a single cut of a circular molecule. In timed digestions, the expectation is for intermediate bands corresponding to (i) multiples of 3.2-kb units and (ii) a 2.6-kb unit plus multiples of 3.2-kb units. Such bands are
readily resolved upon timed HindIII digestion (Fig. 3B, lanes 4 and 6) up to expected sizes of 15.4 and 16.0 kb. Similar results were obtained with EcoRV (Fig. 3C).

**Ratio of trpEG plus ψtrpEG to trpB.** A representative Southern blot used in the determination of the number of copies of trpEG plus ψtrpEG and trpB is presented in Fig. 4. The results indicated that 1 μg of D. noxia-Buchnera DNA contained (3.2 ± 0.8) × 10^8 copies of 3.2 trp plus 3.2 ψtrp and (2.6 ± 0.7) × 10^7 copies of trpB. In these experiments, the level of 2.6 trp could be determined only at the highest D. noxia-Buchnera DNA concentration and corresponded to (0.44 ± 0.11) × 10^8 copies per μg of D. noxia-Buchnera DNA. This value is in agreement with the observations that there are approximately seven copies of the 3.2-kb unit for each of the 2.6-kb units. The total number of copies of 3.2 trp plus 3.2 ψtrp plus 2.6 trp is therefore about 3.6 × 10^8/μg of endosymbiont-aphid DNA. This represents about a 14-fold amplification over the level of trpB, or about 1.8 plasmids of (3.2 trp) plus (2.6 trp) plus (3.2 trp) or 1.8 units of 3.2 trp per Buchnera cell.

**Overview of the nucleotide sequence comparisons.** To facilitate comparisons, we have split the 9,034-bp DNA fragment into three segments, each of which ends in the nucleotide corresponding to the last amino acid of trpG or ψtrpG. A genetic map of each segment is presented in Fig. 5. It should be noted that the fragment designated 3.2 trp is a composite containing the front and terminal portions of 3.2 trp of the 9,034-bp fragment shown in Fig. 1. Comparisons of 3.2 trp and 3.2 ψtrp (Fig. 5) indicate that the regions containing the major differences are restricted to the segment designated ori and the initial 430 nucleotides (nt) of trpE and ψtrpE. Since 3.2 trp and 2.6 trp are similar, having only 13 differences (Fig. 5), most of the subsequent comparisons will involve only 3.2trp and 3.2 ψtrp. A summary of these differences is presented in Table 1.

**Putative ori and promoter.** Buchnera species associated with *Schizaphis graminum, R. padi, R. maidis,* and *A. pisum* have a conserved 511- to 603-nt sequence upstream of trpE, which has features characteristic of an origin of replication (ori) (39, 40). Among these features are a 9-nt sequence known as a DnaA box (consensus TTATCCACA) and regions of low G+C content. The nucleotide sequences conserved in these four *Buchnera* strains and in ori/3.2 trp and ori/3.2 ψtrp are presented in Fig. 6. In all cases, there is conservation of three 9-nt sequences identical or related to DnaA boxes. In four previously studied *Buchnera* strains, the first DnaA box is identical to the consensus. In ori/3.2 trp and ori/3.2 ψtrp, the second DnaA box is identical to the consensus while the other two differ by 1 nt...
An analysis of the differences between ori/3.2trp and ori/3.2c trp is presented in Table 1. Most differences occur following nt 163. Up to this nucleotide, there are only 2 differences, while following it, there are 101 differences. Among the major modifications is a 20-bp insertion into ori/3.2c trp (nt 394) and changes which eliminate a potential ribosome-binding site (nt 511 to 513; deletion at nt 519 to 523). Starting at nt 446 is the sequence TTGCTT... 14 nt ... AGTTAT; by analogy with putative promoters of Buchnera rRNA operons (4), this could be a putative 235 and 210 region preceding 3.2 trpE.

In ori/3.2c trp, this sequence is modified in two positions (C → A, A → G). The changes from ori/3.2trp to ori/3.2c trp after nt 160 (Fig. 6) probably inactivate the 3.2c trpE promoter region as well as the ribosome-binding site that precedes it.

A 9-nt sequence (Fig. 6, nt 73 to 81) is directly repeated 555 bp upstream of both ori/3.2trp and ori/3.2c trp. In 2.6c trp, the 555-bp segment and one of the 9-nt repeats are deleted.

**trpE and trpG.** Of the differences found between 3.2trpE and 3.2c trpE, 85% (or 106 differences) are found in the first 430 nt (Fig. 5; Table 1). The total number of changes causing premature termination of the protein (frameshifts and stop codons) is 14. In comparison with trpE, there is a lower frequency of change in 3.2etrpG to 3.2c trpG (8 versus 1.7%). Two of these changes involve premature termination. A 15-nt intergenic segment separates trpE from trpG and ψtrpE from ψtrpG. These segments are identical except for a 1-nt difference between the sequences in 3.2trp and 3.2c trp.

**Evolutionary analysis.** Figure 7 shows the results of a phylogenetic analysis of trpE and trpG genes and pseudogenes from D. noxia and from two related aphids. In agreement with the above results, the 2.6- and 3.2-kb unit pseudogenes are most similar to each other and more closely related to the corresponding D. noxia gene than to genes of other species. Also in agreement with the above results, the c trpG sequences are less divergent from trpG than c trpE is from trpE, as reflected by the longer branch leading to c trpE.

For both trpE and trpG, the longer branch leading to D. noxia than to A. pisum suggests that substitution rate has been higher in the D. noxia lineage. An increased frequency of substitutions resulting in amino acid replacements is expected if selection on loci involved in tryptophan biosynthesis is relaxed by a reduced need for tryptophan provision by D. noxia-associated Buchnera species.

**DISCUSSION**

**General properties of the plasmids.** The trpEG-containing plasmids of D. noxia-associated Buchnera species consist of

---

**TABLE 1. Number of differences between 3.2trp and 3.2c trp**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>trpG to ori/3.2etrpG to ori</th>
<th>ori/3.2etrpE to ori</th>
<th>3.2etrpE to 3.2c trpE</th>
<th>3.2etrpG to 3.2c trpG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no. (%) of differences</td>
<td>6 (1.1%)</td>
<td>103 (19.7%)</td>
<td>125 (8.0%)</td>
<td>10 (1.7%)</td>
</tr>
<tr>
<td>No. of positions compared</td>
<td>525</td>
<td>524</td>
<td>1,564</td>
<td>595</td>
</tr>
<tr>
<td>No. of inserted nucleotides</td>
<td>1</td>
<td>24</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>No. of deleted nucleotides</td>
<td>0</td>
<td>14</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>No. of frameshifts</td>
<td>9</td>
<td>1</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>No. of stop codons</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>No. of silent substitutions</td>
<td>16</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>No. of amino acid changes</td>
<td>63</td>
<td>8</td>
<td>12</td>
<td>4</td>
</tr>
</tbody>
</table>
tandem repeats of 3.2-kb units and a 2.6-kb unit (Fig. 1). Each plasmid contains a single copy of trpEG and tandem repeats of trpEG pseudogenes. The unit composition of the principal plasmid is $(3.2 \text{ trp})^6$ plus $(3.2 \text{ trp} + 2.6 \text{ c trp})^1$. $3.2 \text{ trp}$ and $2.6 \text{ c trp}$ are adjacent to each other. $2.6 \text{ c trp}$ differs from $3.2 \text{ trp}$, primarily in containing a 555-bp deletion. Two minor plasmids differ from the principal plasmid in containing five and seven tandem repeats of $3.2 \text{ trp}$. It is not known if each aphid or each Buchnera cell has one or a mixture of these plasmids.

The ratio of trpEG plus c trpEG to trpB (single-copy chromosome) is about 14, indicating the presence of about 1.8 plasmids per D. noxia-associated Buchnera cell or 1.8 copies of trpEG per Buchnera cell. The extent of trpEG plus c trpEG amplification relative to trpB is similar to that observed in Schizaphis graminum-associated Buchnera bacteria (21). The major difference is that there is no evidence for trpEG pseudogenes in the plasmid from the latter endosymbiont.

The major nucleotide sequence difference between the units containing trpEG and c trpEG is in the first 430 nt of the portion coding for trpE and the 360 nt directly upstream of trpE (Fig. 5 and 7). In the case of trpE, the changes lead to premature termination of the nascent peptide. In the case of the upstream regions, the changes consist of the removal of a ribosome binding site and, more speculatively, inactivation of the DNA concerned with trpE expression (Fig. 6). All of these changes should result in silencing of trpEG and a major decrease in the amount of TrpEG protein. In contrast to these changes, there is conservation of an upstream 160-nt segment containing three 9-nt sequences identical to or resembling DnaA boxes (39). This 160-nt region also has similarities to a DNA segment in the same position of trpEG-containing plasmids of Buchnera strains from four other species of aphids (Fig. 6) (21, 32). The general characteristics of this region resemble ori (discussed below), and its similarity in the endosymbions of five species of aphids suggests that plasmid-associated trpEG amplification occurred as a single event in an ancestor of these species.

Gene silencing. The major finding of our study which requires explanation is the dichotomy between the retention of tandem repeats on plasmids and the silencing of most of their trpEG genes. This is in contrast to other systems, in which gene amplification occurs as a mechanism of increasing the production of an enzyme activity which is growth limiting (1). Once the stimulus which led to growth limitation is removed, the number of tandem repeats is usually rapidly reduced to a single
copy (1, 6). In *D. noxia*-associated *Buchnera* species, there must be a strong selective pressure for the retention of tandem repeats and the silencing of most of the copies of *trp*EG. We do not have a definitive explanation for our results but do to offer the following speculations.

As previously discussed, *trp*EG amplification in some aphid species is a means by which the endosymbiont overproduces tryptophan for the aphid host (4, 5, 21, 32). Since the aphid requires this amino acid, its overproduction may, in part, account for the short development time of aphids which have plasmid-associated *trp*EG (4, 5). *D. noxia* causes major structural damage to some host plants (such as wheat), resulting in the dissolution of plant cell contents (16) and possibly making available free tryptophan. This aphid may no longer be totally dependent on the synthesis of this amino acid by the endosymbiont when feeding on a susceptible plant. The synthesis by the endosymbiont of *TrpEG* consumes much more energy than the synthesis of *trp*EG DNA (30), and consequently there may be selection against unnecessary production of this protein. If there is selective pressure against reduction of the tandem repeats (discussed below), the synthesis of *TrpEG* could be prevented by inactivation of the structural gene as well as inactivation of the region involved in its expression, as appears to be the case in *D. noxia*-associated *Buchnera* plasmids (Fig. 5). The retention of one *trp*EG copy per plasmid may still allow sufficient tryptophan synthesis by the endosymbiont to support the growth of *D. noxia* on less suitable plants, which may be alternative sources of food required for its survival.

A possible analogy to the present observations are the previous comparative studies of the tryptophan operon involving *Shigella dysenteriae* and *Escherichia coli* (25, 26). Although these two organisms have for a long time been assigned to two different genera, this conclusion is not in agreement with current taxonomic analysis, which indicates their assignment to a single species (33). *E. coli* is a common inhabitant of the intestinal tract of warm-blooded animals and has no growth factor requirements, being able to grow on a minimal medium (33). Its biosynthetic capabilities are probably important for its survival in the intestinal tract in competition with many other organisms. *Shigella dysenteriae* differs from most strains of *E. coli* in that it is able to invade and proliferate within cells of the intestinal tract (33). This habitat is probably nutrient rich, and *Shigella dysenteriae* may not require many of its biosynthetic genes. Consistent with these speculations is the fact that *Shigella dysenteriae* requires many amino acids, including tryptophan (25).

The basis of the requirement for tryptophan in *Shigella dysenteriae* has been extensively studied (25, 26). It has been found that there are four modifications of the tryptophan operon involving 2-nt changes in the promoter region and two changes which inactivate *trp*E. The result of these changes is a major decrease in the amount of *trp* mRNA made by *Shigella dysenteriae* and a major reduction in the amount of *TrpEG*. These results are similar to those obtained in the present work in that the target sites of the changes are the promoter region and the structural gene for *TrpE*. In *Shigella dysenteriae*, natural selection has favored silencing of the *trp* genes rather than deletion of the *trp* operon.

**Retention of ori.** A characteristic of bacterial origins of chromosomal replication is the presence of several DNA boxes as well as regions of low G+C content (39). Bidirectional chromosomal replication is initiated when the DNA protein binds to ori (40). The location of *Buchnera* ori is not known. The gene for *Schizaphis graminum*-associated *Buchnera* DNA has been cloned and sequenced; it resembles the protein from *E. coli* (22). The *E. coli* ori and those of other related bacteria are preceded by a DNA box located in the promoter region of *ori*, and evidence has been presented for autogenous regulation of DNA (24). In *Schizaphis graminum*-associated *Buchnera* species, no DNA boxes were present upstream of *dnaA*, suggesting a different mechanism of regulation of its synthesis (22). Since the increase in the endosymbiont numbers parallels the increase in aphid weight (4), it was speculated that the host controls the rate of DNA synthesis and consequently the initiation of *Buchnera* chromosomal replication (22). The introduction of plasmid DNA containing ori or DNA boxes could be toxic for the host cell, since these sequences are thought to titrate out the DNA protein (17, 28). This mechanism of toxicity is consistent with the present study, in that 3.2*trp* on a high-copy-number vector was toxic to *E. coli* while 2.6*trp*, which lacks the consensus DNA boxes and a DNA box differing in 1 nt from the consensus, was not toxic. Possibly, the aphid ancestor in which *trp*EG became plasmid borne and amplified adjusted its regulatory mechanism governing the level of DNA protein so as to allow for the presence of multiple additional ori regions and DNA boxes. Once this compensatory mechanism was established, the removal of plasmid ori and DNA boxes was toxic to the endosymbiont, since it would have resulted in an excess of DNA and, consequently, in an increase in the initiation of *Buchnera* chromosome replication (35). Since this regulatory mechanism is complex (possibly involving the host), it may have been simpler to retain the initial number of ori-containing tandem repeats and inactivate *trp*EG rather than to decrease the number of tandem repeats by deletion. The consequences of this would be the retention of the same number of DNA boxes but a major reduction in the amount *TrpEG*, as appears to be the case in the *D. noxia*-associated *Buchnera* plasmids.

The high sequence similarity between 3.2*trp* and 2.6*trp* suggests that recombination between the units containing pseudogenes does occur (Fig. 5). Without recombination, the pseudogenes would change independently and have extensive differences from each other. The presence of a 555-bp deletion in 2.6*trp* is also consistent with the occurrence of recombination.

**Evolutionary analysis.** *D. noxia*-associated *Buchnera* *trp*EG genes appear to have undergone accelerated evolution (Fig. 7). Faster evolution is especially apparent in the first 430 nt of *trpE* (Fig. 5). The number of differences between *Buchnera* species from *R. padi*, *R. maidis*, and *Schizaphis graminum* in the same segment of *trpE* DNA ranges from 59 to 70. In this region, there are 106 differences between *D. noxia*-associated *Buchnera* 3.2*trpE* and 3.2*trpE*. This increase in the rate could be due to the acquisition by *D. noxia* of a mechanism which results in extensive damage of plant cells, thereby making available free amino acids and reducing the selective pressure for the retention of a full complement of functional plasmid-associated *trp*EG tandem repeats.

**Overview of trp evolution in aphids.** We would like to summarize the results of our present and past studies (21, 32) by suggesting the following scenario for *trp* evolution. The ancestor of present-day *Buchnera* species had *trp* genes arranged as two linkage groups [*trp*EG and *trp*DC(F)rA] on its chromosome. In the ancestor of the aphids *D. noxia*, *A. pisum*, *Schizaphis graminum*, *R. padi*, and *R. maidis*, there was *trp*EG amplification in the endosymbiont. The overproduction of this amino acid, as well as other amino acids (8), led to a decrease in the development time of the aphid. The evolution of the *Buchnera* plasmid- and chromosome-associated *trp* genes, as well as other chromosomal genes and host genes, was congruent, indicating a lack of genetic exchange between endosymbionts. In *Buchnera* species from *D. noxia*, there was an increase...
in the rate of trpEG evolution and the formation of trpEG pseudogenes as a consequence of the acquisition by this aphid lineage of means to derive tryptophan from the diet, thus reducing the dependence of the aphids on the biosynthetic capability of the endosymbionts.

**ACKNOWLEDGMENTS**

This material is based on work supported by the National Science Foundation under award MCB-9402813 and by the University of California Agricultural Experiment Station.

We thank M. A. Clark for the cloning and sequencing of the D. noxia-associated Buchnera 0.68-kb trpB DNA fragment, B. Beaman for helpful discussions, and S. R. Kushner for the gift of pWKS129/130.

**REFERENCES**