Vertical Transmission of Biosynthetic Plasmids in Aphid Endosymbionts (Buchnera)

JENNIFER J. WERNEGREEN* AND NANCY A. MORAN
Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721

Received 2 February 2000/Accepted 16 October 2000

This study tested for horizontal transfer of plasmids among Buchnera aphidicola strains associated with ecologically and phylogenetically related aphid hosts (Uroleucon species). Phylogenetic congruence of Buchnera plasmid (trpEG and leuABC) and chromosomal (dnaN and trpB) genes supports strictly vertical long-term transmission of plasmids, which persist due to their contributions to host nutrition rather than capacity for infectious transfer. Synonymous divergences indicate elevated mutation on plasmids relative to chromosomal genes.

Bacterial genomes are characterized by remarkable plasticity that allows rapid genetic adaptations to environmental changes (reviewed in references 3, 33, 46). Plasmids, extrachromosomal DNA molecules that replicate autonomously, contribute to this plasticity by mediating lateral gene transfer among bacterial species and genera (15, 17, 21, 40, 57, 59, 65) and even between kingdoms (19, 24). In addition to their role in lateral gene transfer, plasmids also function in gene amplification and overexpression (46, 47). Just as chromosomal duplications are a common mechanism for increasing gene dosage in response to fluctuations in the environment (47, 54), amplification of loci on plasmids may be adaptive when selection favors increased gene dosages (12, 20).

In Buchnera aphidicola, the endosymbiotic symbiont of aphids, genes for the biosynthesis of tryptophan (trpEG) and leucine (leuABCD) often occur on multicopy plasmids (pTrpEG and pLeu, respectively) (5, 6, 10, 31, 48, 49, 52, 63, 64). Comparative sequence analysis indicates that the ancestral location for both trpEG and leuABCD genes was the Buchnera chromosome, not an exogenous plasmid (7, 49, 63). This movement of chromosomal loci onto plasmids is considered a host-beneficial adaptation of Buchnera to overproduce these essential amino acids that are lacking in the hosts’ diet of plant sap.

The role of horizontal transfer in the evolution of Buchnera biosynthetic plasmids remains unclear. In contrast to facultative symbionts such as Rhizobium and Vibrio, lateral gene transfer in Buchnera may be highly constrained since this obligate symbiont spends its entire life cycle within specialized host cells (bacteriocytes) (11, 43). In accordance with this hypothesis, several previous studies show phylogenetic congruence among chromosomal (trpB and 16S rRNA) and plasmid (trpEG and leuABCD) genes of Buchnera associated with the family Aphididae and suggest a lack of plasmid transfer in this symbiont group (5, 6, 10, 22, 48, 49, 51, 63, 64). However, recent work suggests horizontal transfer of the plasmid-encoded repA1 gene in Buchnera of Pemphigus sypotothecae (62).

Most previous studies were based on sampling Buchnera associated with different aphid genera and cannot address the issue of plasmid transfer among closely related strains, which may occur via biological vectors or acquisition of DNA from the environment (60). In order to maximize the chance of detecting gene transfer among related Buchnera lineages, we sampled Buchnera of Uroleucon, a recent radiation of aphids that specialize on Asteraceae and often share host plants, habitats, secondary endosymbionts, and parasitoids (42, 50). We compare phylogenies of chromosomal genes (dnaN and trpB) and plasmid-encoded genes (trpEG and leuABC) to test for plasmid transfer in this symbiont group.

Phylogeny reconstruction. Collection data, aphid DNA extractions, and standard PCR conditions were described previously (42). The PCR was used to amplify three gene regions of Buchnera: dnaN (1,107 bp), leuABC (3,919 bp), and trpEG (1,767 bp) (primer sequences available upon request). DNA sequences were obtained as described previously (42) directly from PCR products or TA clones of PCR fragments. GenBank numbers for sequences obtained here and for previously published sequences are given in Table 1. Translated DNA sequences were aligned by using Megalign (DNAltar).

Genealogies of each of the four gene regions and for combined data were estimated by using maximum parsimony (MP) and maximum likelihood (ML) (Paup* 4 [56]). MP trees were estimated by heuristic searching, and confidence in nodes was assessed by bootstrapping (100 replications). MP trees estimated for the subset of taxa available for each locus (Fig. 1, taxa in bold) agree with relationships shown in the larger bootstrap trees for all available taxa (Fig. 1). These MP trees were generally very similar across genes. ML phylogenies were estimated for the subset of Buchnera lineages sequenced for each gene region after excluding third codon positions. ML parameters and topologies were alternatively estimated until there was no improvement in the likelihood score, according to the successive approximation method suggested by Swofford (55). The proportion of invariant sites and base frequencies were set to empirical levels, and substitution rates were allowed to vary among sites according to a gamma distribution (four site categories) under the Hasegawa-Kishino-Yano model of substitution. Phylogenies and the ML parameters alpha (the gamma shape parameter) and transition/transversion ratio were estimated separately for each region. Only two ML topologies were found: (i) that of leuABC and dnaN and (ii) that of trpB, trpEG, and combined data (Fig. 2).
MP and ML estimates give similar phylogenies for all gene regions. Notably, the MP and ML trees for combined data are identical. Slight discrepancies between MP and ML estimates result primarily from the placement of two taxa, *Buchnera of Uroleucon erigeronense* and *Uroleucon caligatum*. These discrepancies are only weakly supported, as seen in the low MP bootstrap values (Fig. 1) and short internal branches on ML trees (Fig. 2). The relationships at each gene generally agree with relationships among the *Uroleucon* hosts (14).

**Phylogenetic congruence among loci.** Outgroup species (*Rhopalosiphum padi* and *Schizaphis graminum*) were excluded from tests of phylogenetic congruence to avoid biasing the outcome towards congruence. First, we tested the null hypothesis, using TREEMAP (44), that MP trees for each data set are no more congruent than expected by chance (i.e., randomly related). All pairs of MP trees were more similar than expected by chance (P < 0.001 for each comparison). However, disproving the null hypothesis of random relatedness provides only weak evidence for congruence, since gene transfer may not erase all traces of historical associations (see reference 14). We therefore tested the null hypothesis that different gene regions support the same topology. The Kishino-Hasegawa test evaluates whether a data set has a significantly better likelihood score across its own ML tree than across the alternative ML topology (28) (using Paup*4). Similarity in likelihood scores for both ML trees indicates that discrepancies between the two *Buchnera* ML phylogenies are not statistically significant for any gene region (Table 2).

This phylogenetic congruence of plasmid and chromosomal genes strongly supports a lack of plasmid transfer among *Buchnera* strains associated with aphid hosts that share habitats, host plants, and parasitoids and secondary endosymbionts (50). A recent study found congruence of gene genealogies in *Buchnera of Uroleucon ambrosiae*, suggesting strictly vertical transfer even within the same host species (22). These results support previous conclusions of congruence among *Buchnera* genealogies and contribute to the larger picture of vertical plasmid transmission across millions of years (5, 6, 10, 48, 49, 52, 63, 64). Our data suggest that the single proposed instance of plasmid transfer in *Buchnera* may represent a very rare event that occurred early in the evolution of the Pemphigidae (62). This plasmid stability in *Buchnera* contrasts with genome

![Figure 1](http://jb.asm.org)
fluidity in most bacterial species, where plasmids mobilize ecologically important features such as pathogenic and symbiotic capacities (69) and antibiotic resistance (4, 37) and contribute to the mosaic-like genome structure of some bacterial genomes (26, 32, 34, 41). For example, in Escherichia coli, the close free-living relative of Buchnera, incongruence among genealogies of chromosomal and plasmid-encoded genes indicates several recombination and horizontal transfer events (35).

Selection for plasmid maintenance. The maintenance of bacterial plasmids has been attributed to a combination of infectious transfer and selection for plasmid-encoded traits (38). Since the two biosynthetic plasmids of Buchnera experience little if any lateral transfer, they must be maintained solely by selection for plasmid-encoded traits. In endosymbionts, selection may occur within hosts (resulting from differential replication of different endosymbiont genotypes within an individual host) and between hosts (resulting from differential reproductive rates of hosts that contain different symbiont genotypes) (1, 45). At the level of within-host selection, plasmid amplification of biosynthetic genes in Buchnera is probably neutral or deleterious, since the overproduction of tryptophan and leucine and the replication of plasmids may be costly to individual Buchnera cells (36, 53). Any selection favoring plasmid maintenance in Buchnera must occur between aphids, which require symbiotic biosynthetic functions for adequate nutrition. This impact of host-level selection may explain the prevalence of these two plasmids in Buchnera of the Aphididae, in which relatively rapid growth and high fecundity may increase physiological demands for amino acids (7). Host-level selection may also explain the parallel changes in level of amplification of trpEG and leuABCD in particular aphid species (58).

With the above reasoning, selection on bacterial cells will tend to favor plasmid loss while selection on aphid hosts will favor plasmid maintenance. Such conflict may be partially resolved by an attenuation of negative effects of plasmids on bacterial fitness (e.g., see reference 9). Based on the divergence times of aphids with plasmid-bearing Buchnera (about 50 to 70 million years for the family Aphididae) and the estimated generation time of Buchnera (about 50 doublings per
year [13]), pTrpEG and pLeu have been vertically transmitted with the *Buchnera* chromosome for approximately 2.5 to 3.5 billion bacterial generations, over which time selection may have minimized deleterious effects of plasmids on bacterial fitness. One possible mechanism by which individual *Buchnera* cells may benefit from (or be “addicted” to) pTrpEG is through selection to preserve *dnaA* boxes borne by this plasmid (30). These sites may titrate DnaA, a protein that is also involved in initiation of chromosomal replication.

**Elevated divergence at plasmid-borne genes.** Some plasmids may experience elevated rates of sequence evolution compared to chromosomal genes due to higher rates of adaptive fixation or higher mutation rates (18, 61). The latter might arise from more frequent recombination (27), greater densities of transposons (2), dependence on different, more error-prone polymerases (25, 29), or higher rates of transcription (16). The vertical transmission of *Buchnera* plasmids provides a rare opportunity to contrast rates of sequence divergence among completely linked, autonomously replicating DNA molecules. In many bacterial species, selection for codon usage at particular loci (adaptive codon bias) may lead to differences in rates of synonymous substitution among genes (50). However, since *Buchnera* lacks adaptive codon bias (66), synonymous substitution rates are expected to approximate neutral mutation rates. Under the hypothesis that different replicons have equal mutation rates, divergence at synonymous sites is expected to be similar for plasmid and chromosomal genes.

Synonymous divergences were estimated for each gene region across phylogenetically independent pairs of *Buchnera*. Synonymous divergences were estimated using the method of Li (39), adjusting for moderate levels of sequence divergence (using the program Molecular Evolutionary Analysis [E. Moriyama, Yale University]) and the maximum-likelihood-based method of Goldman and Yang (23) (codeML package of PAML [67]). Compared to other methods, this likelihood approach accounts for unequal base (codon) frequencies and biased transition/transversion ratios and provides a more realistic evolutionary model for DNA sequences with extreme

TABLE 3. Pairwise estimates of synonymous divergences at chromosomal (*dnaN* and *trpB*) and plasmid (*leuABC* and *trpEG*) genes for six phylogenetically independent pairs of *Buchnera* isolates, showing generally higher divergences at the plasmid-encoded loci for any given pairwise comparison.

<table>
<thead>
<tr>
<th>Method and comparison*</th>
<th><em>dnaN</em> (Divergence)</th>
<th><em>trpB</em> (Divergence)</th>
<th><em>leuABC</em> (Divergence)</th>
<th><em>trpEG</em> (Divergence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uamb vs Urud</td>
<td>0.201 (0.042)</td>
<td>0.333 (0.081)</td>
<td>0.360 (0.045)</td>
<td>0.314 (0.040)</td>
</tr>
<tr>
<td>Uson vs Uaen</td>
<td>0.393 (0.070)</td>
<td>0.857 (0.199)</td>
<td>1.063 (0.117)</td>
<td>0.910 (0.099)</td>
</tr>
<tr>
<td>Urur vs Uhel</td>
<td>0.499 (0.095)</td>
<td>0.815 (0.171)</td>
<td>0.897 (0.097)</td>
<td>0.799 (0.082)</td>
</tr>
<tr>
<td>Ue vs Ucal</td>
<td>0.578 (0.154)</td>
<td>1.260 (0.277)</td>
<td>1.350 (0.138)</td>
<td>2.688 (0.444)</td>
</tr>
<tr>
<td>Ap vs Mi</td>
<td>1.909 (0.358)</td>
<td>1.735 (0.394)</td>
<td>2.426 (0.336)</td>
<td>2.153 (0.299)</td>
</tr>
<tr>
<td>Rp vs Sg</td>
<td>0.710 (0.148)</td>
<td>0.851 (0.211)</td>
<td>1.012 (0.102)</td>
<td>1.368 (0.283)</td>
</tr>
<tr>
<td>Ks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uamb vs Urud</td>
<td>0.116 (0.029)</td>
<td>0.137 (0.033)</td>
<td>0.134 (0.014)</td>
<td>0.190 (0.025)</td>
</tr>
<tr>
<td>Uson vs Uaen</td>
<td>0.197 (0.038)</td>
<td>0.301 (0.058)</td>
<td>0.340 (0.027)</td>
<td>0.561 (0.059)</td>
</tr>
<tr>
<td>Urur vs Uhel</td>
<td>0.201 (0.034)</td>
<td>0.356 (0.064)</td>
<td>0.312 (0.028)</td>
<td>0.512 (0.049)</td>
</tr>
<tr>
<td>Ue vs Ucal</td>
<td>0.389 (0.065)</td>
<td>0.442 (0.076)</td>
<td>0.902 (0.038)</td>
<td>1.202 (0.135)</td>
</tr>
<tr>
<td>Ap vs Mi</td>
<td>0.679 (0.108)</td>
<td>0.623 (0.110)</td>
<td>0.663 (0.056)</td>
<td>1.447 (0.271)</td>
</tr>
<tr>
<td>Rp vs Sg</td>
<td>0.242 (0.042)</td>
<td>0.295 (0.057)</td>
<td>0.343 (0.027)</td>
<td>0.604 (0.063)</td>
</tr>
</tbody>
</table>

* Divergences were estimated by Goldman and Yang’s dS (23) and Li’s Ks (39), adjusting for moderate sequence divergences. *Buchnera* taxa are listed by the abbreviation of their aphid host names (see Table 1).

FIG. 3. Phylogenetically independent pairwise estimates of synonymous divergence at *Buchnera* *dnaN*, *trpB*, *leuABC*, and *trpEG* genes, using a maximum likelihood-based estimation (67, 68). Error bars indicate the standard errors of individual pairwise estimates. See Table 1 for abbreviations.
base compositions (reviewed in reference 68). Assumptions of the likelihood method as implemented here include a constant base composition and a uniform rate of substitution across codons of a particular gene (the shape parameter alpha = infinity).

Discrepancies between Goldman and Yang's (23) and Li's (39) methods illustrate the effects of accounting for base composition when calculating sequence divergences (Table 3). However, both methods show higher divergences at trpEG and leuABC compared to those at dnaN and trpB for four of the six pairwise comparisons (Fig. 3; Table 3). Higher synonymous divergence at trpEG than at chromosomal genes agrees with previous studies based on smaller, more divergent data sets (13, 48) and suggests elevated neutral mutation rates on pTrpEG. In contrast with the elevated synonymous divergence at leuABC in our data set, previous studies based on fewer and more divergent taxa did not find elevated rates at leucine biosynthesis genes (5, 13). The Uroleucon sample used here consists of numerous recently diverged isolates, provides low standard errors for divergence estimates, and offers improved ability to compare divergence levels among loci. Overall, our analysis suggests that the mutation rates for both pTrpEG and pLeu may show a moderate increase over that of the chromosome, at least in Buchnera of Uroleucon. Mechanisms for higher mutation rates at plasmid loci may include the use of different DNA polymerases that vary in error rate (29) or higher levels of transcription, which may elevate mutation rates (8, 16).

Nucleotide sequence accession numbers. GenBank numbers for sequences obtained in this study and for previously published sequences are given in Table 1.

This work was supported by a National Institutes of Health post-doctoral training grant in Molecular Insect Science (Center for Insect Science, University of Arizona) to J.J.W. and a National Science Foundation grant (DEB-9415413) to N.A.M.

We thank J. Sandstro ¨m for obtaining most of the original Uroleucon samples and P. Baumann for DNA extractions. Three anonymous reviewers gave helpful comments.

REFERENCES

This page contains a list of references, possibly from a scientific or academic journal. The text is formatted in a citation style commonly used in academic writing, providing information about various studies and findings related to bacterial biology, evolution, and symbiosis. Each entry includes the authors' names, publication year, title, and details of the publication, such as journal names, volume, issue, and page numbers. The format is consistent, adhering to a citation style that facilitates accessibility and comprehensibility for researchers and students in the field of molecular biology and microbiology.