Biochemical and Biophysical Research Communications **292,** 263–267 (2002)
doi:10.1006/bbrc.2002.6646, available online at http://www.idealibrary.com on IDEAL®

An Experimental Validation of Orphan Genes of *Buchnera,* a Symbiont of Aphids

Sayaka Shimomura, Shuji Shigenobu,[1] Mizue Morioka, and Hajime Ishikawa[2]
*Department of Biological Sciences, Graduate School of Science, University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

Received February 12, 2002

Although *Buchnera* sp. APS, an intracellular symbiont of pea aphids, is a close relative of *Escherichia coli,* its genome has been extensively modified because of its prolonged intracellular life. In our previous studies on the *Buchnera* genome, computer analysis predicted three "orphan" genes, *yba2, yba3,* and *yba4,* which are open reading frames (ORFs) with no homologs in the database. In this paper, we successfully validated all these orphan genes by RT-PCR and Northern hybridization. The present study also revealed that *yba3* and *yba4* formed an operon, suggesting that they function in concert. Sequences around transcriptional start sites suggests that these genes are under the control of sigma 70. In view of codon usage and AT bias observed in these genes, it is likely that *Buchnera* have maintained them for an evolutionarily long time.

***Key Words:* *Buchnera* sp. APS; bacterial symbiont; orphan genes; transcription.**

Most, if not all, aphids harbor intracellular bacterial symbionts called *Buchnera* in their bacteriocytes, huge cells specialized for this purpose (1, 2). *Buchnera* and their host aphids are intimately mutualistic and indispensable to each other for the growth and reproduction (1, 3, 4–6). Evidence suggests that this symbiosis started 200–250 million years ago (7), and that, since then, *Buchnera* have coevolved with aphids through vertical transmission through generations of the host (8). Our previous study on the complete genome sequence of *Buchnera* sp. APS revealed that this bacterium has undergone massive gene loss, and that its gene repertory well reflects a long history of mutualistic life with the host aphids (9). In the *Buchnera* genome, we identified 583 ORFs, and similarity search permitted us the functional assignment of 500 of them. Other 79 ORFs were homologous to hypothetical genes detected on the genomes of other bacteria. Among the 4 ORFs unannotated, *yba1* turned out to be homologous to part of the *fliK* gene of *E. coli* (our unpublished data, 10). Thus, only 3 ORFs, *yba2, yba3,* and *yba4,* on the *Buchnera* genome have no similarity to any other reported sequence. This is a surprisingly small number, because in all the genomes sequenced to date, at least, 20% of genes are unknown, or orphan, ones. In addition, as many as 569 ORFs in the *Buchnera* genome have orthologs in *E. coli* (3, 9), and in general, the most similar counterparts of these ORFs are those of *E. coli.* These results suggest that the *Buchnera* genome is a subset of the *E. coli* genome, which all the more spotlights the 3 orphan genes in *Buchnera* described above. If they are really active genes, their products can serve as key molecules that characterize *Buchnera,* an obligately intracellular bacterium.

Since the 3 orphan genes were only theoretically predicted *in silico,* in the present study we first validate them experimentally by RT-PCR and Northern hybridization. Subsequently, we analyze the sequences these genes to gain an insight into their possible function and evolutionary origin.

MATERIALS AND METHODS

Materials. A long established parthenogenetic clone of pea aphids, *Acyrthosiphon pisum* (Harris), was maintained on young broad bean plants, *Vicia faba* L. at 16°C in a long-day regime of 16 hours light and 8 hours dark. Total DNA and RNA were extracted from whole body of aphids with DNAzol (GIBCO BRL) and TRIsol reagent (GIBCO BRL). Extracted RNA was treated with DNase I.

RT-PCR. Three sets of primers were constructed, according to the genome sequence of *Buchnera,* to generate specific fragments for *yba2, yba3,* and *yba4* (Fig. 1): yba2, 5′-ATGACAGATGAAGAAAAAAATTTAATAG-3′ and 5′-AAGTTGTCATTGTCAATATCGTCTG-3′; *yba3,* YBA3a, 5′-ATTCCTGAACAACACCATTCA-3′ and YBA3b, 5′-GGTGATGCATTTGGAGGAAG-3′; *yba4,* YBA4a, 5′-ATGTTACACAATTTATTATTATTAAATCCC-3′ and YBA4b, 5′-GTGTATTTTGAATAATTTCATTATCAC-3′. Using total RNA as templates, cDNAs were synthesized with random hexamers and SUPERSCRIPT II reverse transcriptase (GIBCO BRL). The cycling parameters were as follows:

[1] Present address: Center for Integrative Bioscience, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji, Okazaki 444-8585, Japan. Fax: 81-564-55-7581. E-mail: shige@nibb.ac.jp.

[2] To whom correspondence should be addressed at The University of the Air, 2-11 Wakaba, Mihama-ku, Chiba 261-8586, Japan. Fax: 81-43-4186. E-mail: iskw@u-air.ac.jp.

0006-291X/02 $35.00

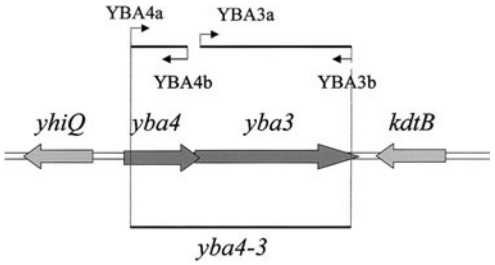

FIG. 1. Arrangement of *yba3* and *yba4*, and primers used for RT-PCR. Coding regions of *yba3*, *yba4*, and surrounding genes are shown by thick arrows indicating the directions of their transcription. Solid lines above the arrows indicate PCR products amplified using primers designated at the both ends. The lowermost line represents the *yba4-3* transcript amplified by RT-PCR using YBA4a and YBA3b as primers.

96°C for 3 min, followed by 94°C for 30s, 58°C for 30s, 68°C for 1 min for 25 cycles and 72°C for 10 min.

Northern hybridization. Samples containing RNA were separated by electrophoresis on agarose gel containing formaldehyde as a denaturant, and transferred onto Hybond N+ membranes (Amersham Pharmacia). Probes specific to *yba2*, *yba3*, and *yba4* were synthesized with the primers used for RT-PCR, and labeled with ^{32}P by random priming using BcaBEST (TaKaRa).

Computer analyses. The nucleotide and protein sequences of *Buchnera* genes were referred to the database of *Buchnera* (http://buchnera.gsc.riken.go.jp/). Similarity searches were performed using BLAST (11, 12), FASTA (13, 14), and SSEARCH (13, 14) against nonredundant protein sequence database (Release 2001-8-16, NCBI). Motif and profile search were performed at the network service of GenomeNet (http://motif.genome.ad.jp/). Amino acid sequence alignments were obtained through the clustalW program (15). Orthology of homologs was examined as described previously (16). Codon Adaptation Index (CAI) analysis was performed using the CAI program, one of the components of the EMBOSS suite (http://www.uk.embnet.org/Software/EMBOSS/).

5′ RACE. To estimate the transcriptional start sites of the *yba2* and *yba4-3* operon, we first performed RT-PCR with several primers constructed for the upstream regions of the translational initiation codons of *yba2* and *yba4*, and predicted approximate start sites. Then, we performed 5′ RACE to analyze in detail these results, using specific primers (SP1) that were constructed opposite to mRNA of *yba2* (5′-AAGTTGTCATTGTCAATATCGTCTG-3′), and *yba4-yba3* (5′-GGTGATGCATTTGGAGGAAG-3′). With these primers, first strand cDNA was extended toward the 5′ end of the transcripts of these genes using total RNA. Subsequently, terminal transferase was used to add a homopolymeric A-tail to the 3′ end of cDNA. Using oligo-dT anchor primer and nested primers (SP2) that were constructed for the inside of cDNA of *yba2* (5′-GAACTGGAATAGTTAGAGGG-3′), and *yba4-yba3* (5′-GTGTATTTTGAATAATTTCATTATCAC-3′), the transcriptional start regions of *yba2*, *yba4-yba3* were amplified by PCR. Throughout these processes, we used 5′/3′ RACE Kit (ROCHE). Sequences of these products were determined using an ABI prism 310 sequencer.

RESULTS

Identification of the Transcripts of yba2, yba3, and yba4

yba2, *yba3*, and *yba4* genes of *Buchnera* were originally identified by the GeneHacker program, a system for gene structure prediction using a hidden Markov Model (HMM) (9, 17). In an effort to confirm the expression of these orphan genes, we performed RT-PCR using the primer pairs constructed according to the sequences around the start and last codons of these genes, and total RNA as templates. As a result, amplicons were detected for all of the three genes as a single band on each lane (Fig. 2). These bands were not detected in the absence of reverse transcriptase. The sizes of these products coincided with those expected from the sequence data (*yba2*, 710 bp; *yba3*, 1073 bp; *yba4*, 319 bp). This result successfully validated the transcription of the three hypothetical genes in the *Buchnera* genome.

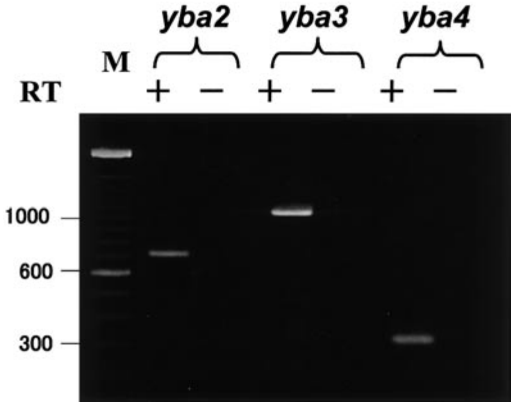

FIG. 2. Expression of orphan genes. Transcripts of *yba2*, *yba3*, and *yba4* were amplified by RT-PCR using specific primers and electrophoresed on agarose gel. (+) PCR after reverse transcription (RT); (−) PCR without RT. M, molecular marker of 100 bp ladder.

Structure of yba2 and yba4-3 Transcripts

To examine the primary structures of the transcripts from the three *Buchnera* genes, we analyzed aphid total RNA by Northern hybridization using *yba2*-, *yba3*-, and *yba4*-specific probes (Fig. 3). The results

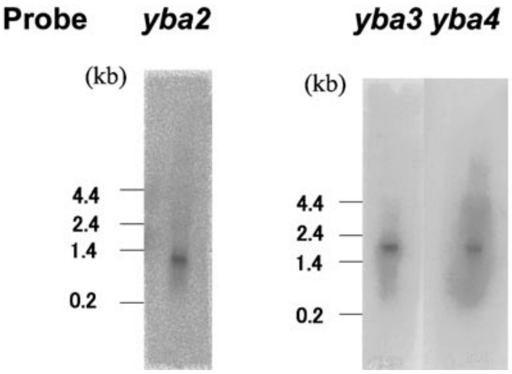

FIG. 3. Northern hybridization of orphan gene products. Total RNA from the whole body of aphids was subjected to agarose gel electrophoresis, hybridized with ^{32}P-labeled probes specific to *yba2*, *yba3*, and *yba4*, and autoradiographed.

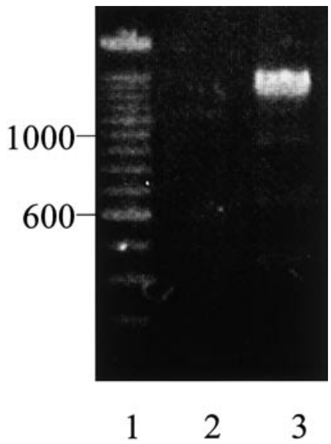

FIG. 4. Identification of *yba4-3* transcript by RT-PCR. The *yba4-3* transcript was amplified by RT-PCR using YBA4a and YBA3b as primers. 1, Molecular marker; 2, PCR without RT; 3, PCR after RT.

further confirmed the expression of these genes. The transcript of *yba2* was about 1 kb, a little longer than the size of its ORF (717 bp). Judging from the orientation of neighboring genes, which is opposite to that of *yba2,* it was likely that the *yba2* gene was transcribed as a monocistronic mRNA.

The *yba3*-specific probe hybridized with the RNA whose length was approximately 2 kb. The *yba4*-specific probe also hybridized with the RNA of the same size (Fig. 3). *yba4* resides just upstream of *yba3,* and the segment extending from the start codon of *yba4* to the termination codon of *yba3* is slightly shorter in length than 2 kb. Taken together, these results suggest that the two genes are expressed as a single operon.

To confirm this, we performed RT-PCR using a primer pair, one was designed for the translational start site of *yba4,* and the other for the termination site of *yba3.* As shown in Fig. 4, the length of the PCR product was about 1.4 kb, suggesting that the two genes were transcribed as a single unit.

Transcriptional Start Sites of yba2 and yba4-3

We performed 5′ RACE with primers complementary to the segments flanking the 5′ end regions of *yba2* and *yba4.* As a result, the transcriptional start sites of these operons were located at 89 bp and 86 bp upstream of their translational start codons, respectively (Fig. 5). As indicated in Fig. 5, there was the ribosome-binding site, or the Shine-Dalgarno sequence, upstream of the translational start codon of each operon. Potential −10 and −35 sequence were also identified upstream of the transcriptional start site of each operon. These sequences and their arrangement were similar to those of the *E. coli* s^{70} promoter consensus sequence.

Codon Adaptation Index

The GC-contents of *yba2, yba3,* and *yba4* were 26.1%, 25.4%, and 20.6%, respectively. These values are nearly the same as the overall GC-content (26.3%) of the whole genome. Then, we performed Codon Adaptation Index (CAI) analysis to assess bias in synonymous codon usage. The scores of *yba2, yba3,* and *yba4* were 0.710, 0.706, and 0.663, respectively, which were comparable to those of other *Buchnera* genes (0.72 on average).

DISCUSSION

In every genome project, computer analysis points out a large number of ORFs, but, in many cases, unable to predict their biological functions. Consequently, in many genomes sequenced to date, many genes have been left as unknown, or orphan, genes even without verifying their expression in organisms. Unlike in many other genomes sequenced to date, orphan genes in the *Buchnera* genome that were predicted by computer analysis are only three (9). In view of the fact that almost all the other *Buchnera* genes have their orthologs in *E. coli,* the closest relative known (3, 18), these orphan genes can be top-rated candidates that characterize the unique lifestyle of this bacterium. Keeping this in mind, in the present study we examined whether or not these orphan genes were actually expressed in *Buchnera* cells. As a result, RT-PCR and Northern hybridization clearly evidenced that *Buchnera* contained transcripts of these genes (Figs. 2–4). In accordance with the result that *yba2* is transcribed, our ongoing proteome analysis using mass spectrometry already allowed us to identify the Yba2 protein in the *Buchnera* cell (our unpublished data).

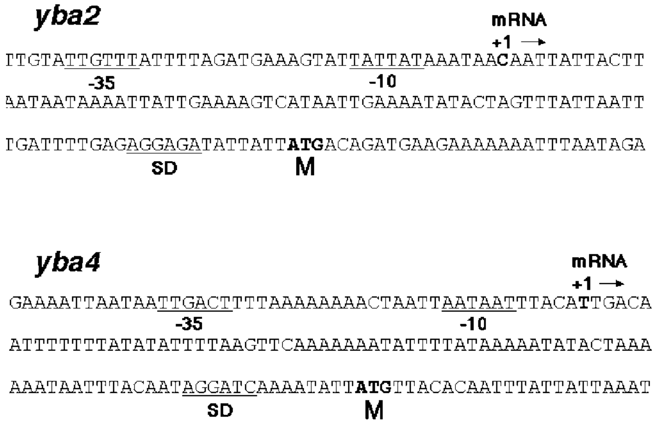
yba2
mRNA
+1 →
TTGTATTGTTTATTTTAGATGAAAGTATTATTATAAATAACAATTATTACTT
-35
-10
AATAATAAAATTATTGAAAAGTCATAATTGAAAATATACTAGTTTATTAATT
TGATTTTGAGAGGAGATATTATTATGACAGATGAAGAAAAAAATTTAATAGA
SD
M
yba4
mRNA
+1 →
GAAAATTAATAATTGACTTTTAAAAAAAACTAATTAATAATTTACATTGACA
-35
-10
ATTTTTTTATATATTTTAAGTTCAAAAAAATATTTTATAAAAATATACTAAA
AAATAATTTACAATAGGATCAAAATATTATGTTACACAATTTATTATTAAAT
SD
M

FIG. 5. Upstream sequences of the *yba2* and *yba4-3* operon. Presumed −10 and −35 regions as well as the Shine-Dalgarno (SD) sequence are shown. +1, start sites of transcription determined by 5′ RACE; M, first amino acids of presumed translational products. Promoter consensus elements, −10 and −35 sequences, are underlined.

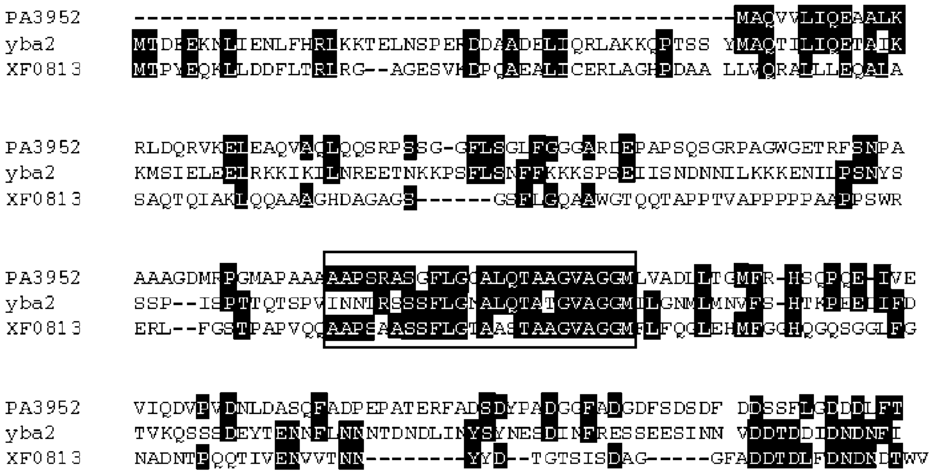
PA3952 ---MAQVVLIQEAALK
yba2 MTDEEKNLIENLFHRLKKTELNSPERDDAADELIQRLAKKQPTSS YMAQTILIQETAIK
XF0813 MTPYEQKLLDDFLTRLRG--AGESVKDPQAEALICERLAGHPDAA LLVQRALLLEQALA
PA3952 RLDQRVKELEAQVAQLQQSRPSSG-GFLSGLFGGGARLEPAPSQSGRPAGWGETRFSNPA
yba2 KMSIELEELRKKIKILNREETNKKPSFLSNFFKKKSPSEIISNDNNILKKKENILPSNYS
XF0813 SAQTQIAKLQQAAAGHDAGAGS------GSFLGQAAWGTQQTAPPTVAPPPPPAAPPSWR
PA3952 AAAGDMRPGMAPAAAAAPSRASGFLGQALQTAAGVAGGMLVADLLTGMFR-HSQPQE-IVE
yba2 SSP--ISPTTQTSPVINNIRSSSFLGNALQTATGVAGGMILGNMLMNVFS-HTKPEELIFD
XF0813 ERL--FGSTPAPVQQAAPSAASSFLGTAASTAAGVAGGMFLFQGLEHMFGGHQGQSGGLFG
PA3952 VIQDVPVDNLDASQFADPEPATERFADSDYPADGGFADGDFSDSDF DDSSFLGDDDLFT
yba2 TVKQSSSDEYTENNFLNNNTDNDLINYSYNESDINFRESSEESINN VDDTDDIDNDNFI
XF0813 NADNTPQQTIVENVVTNN---------YYD--TGTSISDAG-----GFADDTDLFDNDNDTWV

FIG. 6. Alignment of *yba2* homologs. Amino acid sequences predicted from nucleotide sequences of PA3952 of *P. aeruginosa, yba2,* and XF0813 of *X. fastidiosa* were compared. Alignment was achieved using the Internet program Cluster W. A box represents a conserved motif.

After our paper on the *Buchnera* genome had been accepted for publication, genomic sequences of two pathogenic bacteria were published. According to them, two hypothetical genes, PA3952 of *Pseudomonas aeruginosa* (19) and XF0813 of *Xylella fastidiosa* exhibited weak sequence identity with *yba2* (20). In addition, multiple alignment of these genes detected a conserved sequence motif (Fig. 6), suggesting that these gene products would share a common function in *Buchnera* and the two pathogens, though the expression of PA3952 and XF0813 is yet to be validated. Motif and profile searches using this motif as a probe against non-redundant protein data base detected no other protein with a similar motif. Judging from the fact that these three species belong to γ-Proteobacteria (21), it is likely that a *yba2* ortholog was present in the last common ancestor of g-Proteobacteria, followed by gene loss in many lineages including *E. coli* and *Haemophilis influenzae.*

Despite recent accumulation of sequence data, BLAST, FASTA and SSEARCH still failed to detect genes that exhibit significant sequence identity with either *yba3* or *yba4.* One plausible explanation is that these genes have undergone too rapid evolution (22) to identify their orthologs in other organisms. However, this does not necessarily mean that they are inactivated because we successfully identified their bicistronic transcript (Figs. 2–4). It is even probable that these genes have acquired a new function distinct from their ancestral ones, which may be essential for *Buchnera* to maintain the intracellular niche. At present, it is unclear if *yba4* and *yba3* mRNA in the bicistronic transcript are translated independently. As shown in the genome sequence, *yba4* and *yba3* overlap each other by 8 bases, and the presumed initiation codon of *yba3* (ATT) is not the canonical one, which is not apparently preceded by the Shine-Dalgarno sequence. On the other hand, *yba4* contains a stop codon in frame. In addition, a presumed stop codon of *yba3* is not in frame with the initiation codon of *yba4,* but with that of *yba3* (9).

Both GC content and CAI analysis suggest that these orphan genes of *Buchnera* are not recent immigrants from other organisms by lateral gene transfer, but have been retained for an evolutionarily long time. It is likely that ancestral genes of *yba4-yba3,* as well as that of *yba2,* were present in the last common ancestor of g-Proteobacteria, and lost afterward in many lineages. It is unlikely that *Buchnera* retained these genes only by chance because it has lost many genes including even those indispensable to maintain the cell's own identity (23, 24). A probable scenario is that in the course of evolution these genes happened to change so as to function in combination with some genes of host organisms, which, on one hand, prevented them from being lost, and, on the other hand, ensured the symbiotic relationship of *Buchnera* with the hosts. In view of the unique tendency observed with the *Buchnera* genes, this will not necessarily be a far-fetched assumption.

ACKNOWLEDGMENTS

This work was supported by a grant from the Program for Promotion of Basic Research Activities for Innovation Biosciences (ProBRAIN), and Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Science, Sports and Culture. We thank S. Takahashi and S. Masui for their valuable advice and helpful suggestions during the experiments.

REFERENCES

1. Buchner, P. (1965) Endosymbiosis of Animals with Plant Microorganisms, Interscience Publishers, New York.
2. Munson, M. A., Baumann, P., and Kinsey, M. G. (1991) *Buch-*

nera, new genus and *Buchnera aphidicola,* new species, a taxon consisting of the mycetocyte-associated primary endosymbionts of aphids. *Int. J. Syst. Bacteriol.* **41,** 566–568.

3. Baumann, P., Baumann, L., Lai, C. Y., Rouhbakhsh, D., Moran, N. A., and Clark, M. A. (1995) Genetics, physiology, and evolutionary relationships of the genus *Buchnera:* Intracellular symbionts of aphids. *Annu. Rev. Microbiol.* **49,** 55–94.
4. Douglas, A. E. (1998) Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera. Annu. Rev. Entomol.* **43,** 17–37.
5. Houk, E. J., and Griffiths, G. W. (1980) Intracellular symbiotes of the Homoptera. *Annu. Rev. Entomol.* **25,** 161–187.
6. Ishikawa, H. (1989) Biochemical and molecular aspects of endosymbiosis of insects. *Int. Rev. Cytol.* **116,** 1–45.
7. Moran, N. A., Munson, M. A., Baumann, P., and Ishikawa, H. (1993) A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc. R. Soc. Lond. B* **253,** 167–171.
8. Moran, N. A., and Baumann, P. (1994) Phylogenetics of cytoplasmically inherited microorganisms of arthropods. *Trends Ecol. Evol.* **9,** 15–20.
9. Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000) Genome sequence of the endocellular bacteria symbiont of aphids *Buchnera* sp. APS. *Nature* **407,** 81–86.
10. Kutsukake, K., Ohya, Y., and Iino, T. (1990) Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium. J. Bacteriol.* **172,** 741–747.
11. Altshul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410.
12. Altshul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402.
13. Lipman, D. J., and Pearson, W. R. (1985) Rapid and sensitive protein similarity searches. *Science* **227,** 1435–1441.
14. Pearson, W. R., and Lipman, D. J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.
15. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4780.
16. Watanabe, H., Mori, H., Itoh, T., and Gojobori, T. (1997) Genome plasticity as a paradigm of eubacteria evolution. *J. Mol. Evol.* **44** (Suppl. 1), S57–S64.
17. Yada, T., and Hirosawa, M. (1996) Detection of short protein coding regions within the cyanobacterium genome: Application of the hidden Marcov model. *DNA Res.* **3,** 355–361.
18. Clark, M. A., Moran, N. A., and Baumann, P. (1999) Sequence evolution in bacterial endosymbionts having extreme base compositions. *Mol. Biol. Evol.* **16,** 1586–1598.
19. Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrener, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G. K., Wu, Z., and Paulsen, I. T. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* **406,** 959–964.
20. Simpson, A. J., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R., Bueno, M. R., Camargo, A. A., Camargo, L. E., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J., Ferreira, V. C., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C., Lemos, E. G., Lemos, M. V., Lopes, S. A., Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A. M., Madeira, H. M., Marino, C. L., Marques, M. V., Martins, E. A., Martins, E. M., Matsukuma, A. Y., Menck, C. F., Miracca, E. C., Miyaki, C. Y., Monteriro-Vitorello, C. B., Moon, D. H., Nagai, M. A., Nascimento, A. L., Netto, L. E., Nhani, A. Jr., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A., Pereira, H. A. Jr., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., de M Rosa, A. J., de Rosa, V. E. Jr., de Sa, R. G., Santelli, R. V., Sawasaki, H. E., da Silva, A. C., da Silva, A. M., da Silva, F. R., da Silva, W. A. Jr., da Silveira, J. F., Silvestri, M. L., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys, M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., and Setubal, J. C. (2000) The genome sequence of the plant pathogen *Xylella fastidiosa. Nature* **406,** 151–157.
21. Preston, G. M., Haubold, B., and Rainey, P. B. (1998) Bacterial genomics and adaptation to life on plants: Implications for the evolution of pathogenicity and symbiosis. *Curr. Opin. Microbiol.* **1,** 589–597.
22. Shigenobu, S., Watanabe, H., Sakaki, Y., and Ishikawa, H. (2001) Accumulation of species-specific amino acid replacements that cause loss of particular protein functions in *Buchnera,* an endocellular bactrial symbiont. *J. Mol. Evol.* **53,** 377–386.
23. Ishikawa, H. (2001) Symbiotic microorganisms in aphids (*Homoptera, Insecta*): A secret of one thriving insect group. *Korean J. Biol. Sci.* **5,** 163–177.
24. Ochman, H., and Moran, N. A. (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science* **292,** 1096–1099.