



HAL
open science

Gene size reduction in the bacterial aphid endosymbiont, *Buchnera*.

H. Charles, Dominique Mouchiroud, J.R. Lobry, I. Gonçalves, Yvan Rahbé

► **To cite this version:**

H. Charles, Dominique Mouchiroud, J.R. Lobry, I. Gonçalves, Yvan Rahbé. Gene size reduction in the bacterial aphid endosymbiont, *Buchnera*.. *Molecular Biology and Evolution*, 1999, 16 (12), pp.1820-1822. 10.1093/oxfordjournals.molbev.a026096 . hal-00402189

HAL Id: hal-00402189

<https://hal.science/hal-00402189v1>

Submitted on 26 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Letter to the Editor

Gene Size Reduction in the Bacterial Aphid Endosymbiont, *Buchnera*

Hubert Charles,* Dominique Mouchiroud,† Jean Lobry,† Isabelle Gonçalves,† and Yvan Rahbe*

*Laboratoire de Biologie Appliquée, Institut National des Sciences Appliquées 406, Unité Associée Institut National de la Recherche Agronomique 203, Villeurbanne, France; and †Laboratoire de Biométrie, Biologie Evolutive, Unité Mixte de Recherche Centre National de la Recherche Scientifique 5558, Université Claude Bernard Lyon I, Villeurbanne, France

Aphids that feed solely on phloem sap, a diet poor in nitrogenous compounds, harbor intracellular bacteria of the genus *Buchnera* in specialized cells, called bacteriocytes, within their body cavities (Buchner 1965). *Buchnera* are members of the *Enterobacteriaceae* (Munson, Baumann, and Kinsey 1991), and they are transmitted to successive host insect generations via oocyte reinfection (Buchner 1965). Congruent host/bacterium phylogenies suggest that this symbiotic association developed 200–250 MYA (Moran and Baumann 1994), and it is now so close that neither partner could survive independently (Houk and Griffith 1980).

The work of Moran (1996) and Brynne et al. (1998) demonstrated that several *Buchnera* genes exhibit accelerated rates of sequence evolution and comparatively low ratio of synonymous-to-nonsynonymous substitutions for the coding sequences. These findings were attributed to an increased rate of nonsynonymous substitutions, resulting from the population dynamics of *Buchnera* combined with a mutational bias toward A+T. *Buchnera* populations are isolated within the bacteriocytes and experience recurrent bottlenecks during transmission from one host generation to the next, leading to low recombination frequencies between individuals (Moran 1996).

The genome of *Buchnera* contains about 30% G+C (Ishikawa 1987). Most intracellular symbiotic or parasitic bacteria seem to accumulate AT bases in their genomes (reviewed in Heddi et al. 1998). The most striking example is perhaps that of *Mycobacterium leprae*. This bacterium is the only obligate intracellular species of the GC-rich *Mycobacterium* group, characterized by 16S rDNA gene sequences between 56.5% and 59.6% GC, and it is also the most AT-rich species of the genus (56.5%). AT accumulation in obligate intracellular symbiotic or parasitic bacteria could result from an AT mutational bias in conjunction (or not) with the relaxation of selective pressure due to intracellular habitat (Heddi et al. 1998).

The *Buchnera* genome is extremely small, 657 kb compared with 4–5 Mb in related free-living *Enterobacteriaceae* (Charles and Ishikawa 1999). This seems to be a common characteristic of many intracellular symbiotic bacteria, like the endosymbionts of the para-

mecium or the weevil (Soldo, Brickson, and Larin 1983; Charles et al. 1997). Parasitic bacteria also share small genomes of about 1–2 Mb (Herdman 1985). This finding was recently confirmed by genome sequencing projects (reviewed at the following url: www.tigr.org). The stable, protected intracellular environment with no direct interspecies competition may abolish selection pressure on many of the genes implicated in highly integrated and regulated metabolic pathways, some of which are subsequently deleted (Maniloff 1996; Mushegian and Koonin 1996; Razin 1997). Bacteria tend to lose genetic information corresponding to disused metabolic pathways, probably because DNA accumulation slows the population growth rate (Stouthamer and Kooijman 1993). However, it is still not clear whether the genome shrinking also affects coding sequence length or whether or not AT bias is linked to the deletion process.

Deletion mechanisms were analyzed mostly in eukaryotes. Petrov and Hartl (1998) have suggested that genome reduction in some *Drosophila* species may be due to rampant deletions of DNA in regions which are not subject to selective pressure. Other authors assume that organisms with smaller genomes have shorter introns (Duret, Mouchiroud, and Gautier 1995; Hughes and Hughes 1995). Oliver and Marin (1996) suggested that gene length evolution in bacteria could be driven by base composition. Stop codons are AT-rich (i.e., TAG, TAA, and TGA), so stop codon density in AT-rich organisms may be higher than that in GC-rich organisms. For instance, *Escherichia coli* genes were found to be longer than those of *Haemophilus influenzae*, and the shortest genes within each species (*E. coli*, *Bacillus subtilis*, and *H. influenzae*) have the highest AT content (Oliver and Marin 1996).

In this work, we analyzed the deletion mechanisms within the *Buchnera* genome using a sample of 85 protein genes (19 were partial sequences) and their corresponding orthologs in *E. coli* (4.64 Mb, 52% GC) and *H. influenzae* (1.83 Mb, 38% GC), extracted from the GenBank database (May 1, 1998). The *Buchnera* sequences are those of *Buchnera aphidicola* from the aphid *Schizaphis graminum* and of *Buchnera* sp. from *Acyrtosiphon pisum*. The difference between the sequences of these two *Buchnera* species (2.3% divergence in the 16S rDNA gene sequence) was assumed to be small compared with the divergence between *Buchnera* and *E. coli* or *H. influenzae* (9.5% and 14.9% divergence, respectively).

Direct comparison of the complete gene sequences of *Buchnera* and the corresponding orthologous sequences in *E. coli* showed that the *Buchnera* genes were significantly smaller (Wilcoxon signed-ranks test, $P = 0.003$). The mean difference (d) was 8.5 ± 3.3 bp per

Key words: *Buchnera aphidicola*, gene size, GC content, intracellular symbiosis.

Address for correspondence and reprints: Hubert Charles, Laboratoire de Biologie Appliquée, Institut National des Sciences Appliquées 406, Unité Associée Institut National de la Recherche Agronomique 203, 20, Avenue Albert Einstein, 69621 Villeurbanne cedex, France. E-mail: hcharles@insa.insa-lyon.fr

Mol. Biol. Evol. 16(12):1820–1822. 1999

© 1999 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

- BUCHNER, P. 1965. Endosymbiosis of animals with plant microorganisms. Interscience Publishers, New York.
- CHARLES, H., G. CONDEMINE, C. NARDON, and P. NARDON. 1997. Genome size characterization of the principal endocellular symbiotic bacteria of the weevil *Sitophilus oryzae*, using pulse field gel electrophoresis. *Insect Biochem. Mol. Biol.* **27**:345–350.
- CHARLES, H., and H. ISHIKAWA. 1999. Physical and genetic map of the genome of *Buchnera*, the primary endosymbiont of the pea aphid *Acyrtosiphon pisum*. *J. Mol. Evol.* **48**:142–150.
- DURET, L., D. MOUCHIROUD, and C. GAUTIER. 1995. Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. *J. Mol. Evol.* **40**:308–317.
- GONÇALVES, I., M., ROBINSON, G., PERRIERE, and D. MOUCHIROUD. 1999. JaDis: a Java application to compute distance between nucleic acid sequences. *Bioinformatics* **15**:424–425.
- GRAUR, D., Y. SHUALI, and W. H. LI. 1989. Deletions in processed pseudogenes accumulate faster in rodents than humans. *J. Mol. Evol.* **28**:279–285.
- GU, X., and W. H. LI. 1995. The size distribution of insertions and deletions in human and rodent pseudogenes suggest the logarithmic gap penalty for sequence alignment. *J. Mol. Evol.* **40**:464–473.
- HEDDI, A., H. CHARLES, C. KHATCHADOURIAN, G. BONNOT, and P. NARDON. 1998. Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G+C composition of an endocytobiote DNA. *J. Mol. Evol.* **47**:52–61.
- HERDMAN, M. 1985. The evolution of bacterial genomes. Pp. 37–68 in T. CAVALIER-SMITH, ed. *The evolution of genome size*. John Wiley and Sons, New York.
- HOUK, E. J., and G. W. GRIFFITH. 1980. Intracellular symbiotes of Homoptera. *Annu. Rev. Entomol.* **25**:161–187.
- HUGHES, A. L., and M. K. HUGHES. 1995. Small genome for better flyers. *Nature* **377**:391.
- ISHIKAWA, H. 1987. Nucleotide composition and kinetic complexity of the genomic DNA of an intracellular symbiont in the pea aphid *Acyrtosiphon pisum*. *J. Mol. Evol.* **24**:205–211.
- MANILOFF, J. 1996. The minimal cell genome: “on being the right size” *Proc. Natl. Acad. Sci. USA* **93**:10004–10006.
- MORAN, N. A. 1996. Accelerated evolution and Muller’s ratchet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA* **48**:644–649.
- MORAN, N. A., and P. BAUMANN. 1994. Phylogenetics of cytoplasmically inherited microorganisms of arthropods. *Trends Ecol. Evol.* **9**:15–20.
- MUNSON, M. A., P. BAUMANN, and M. G. KINSEY. 1991. *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *Int. J. Syst. Bacteriol.* **41**:566–568.
- MUSHEGIAN, A. R., and E. V. KOONIN. 1996. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. USA* **93**:10268–10273.
- OLIVER, J. L., and A. MARIN. 1996. A relationship between GC content and coding-sequence length. *J. Mol. Evol.* **43**:216–223.
- PETROV, D. A., and D. L. HARTL. 1998. High rate of DNA loss in the *D. melanogaster* and *D. virilis* species groups. *Mol. Biol. Evol.* **15**:293–302.
- RAZIN, S. 1997. Comparative genomics of *Mycoplasmas*. *Wien Klin. Wochenschr.* **109**:551–556.
- SOLDO, A. T., S. A. BRICKSON, and F. LARIN. 1983. The size and structure of the genome of symbiont xenosome particle in the *Parauronema acutum*. *J. Gen. Microbiol.* **129**:1317–1325.
- STOUTHAMER, A. H., and S. A. KOIJMAN. 1993. Why it pays for bacteria to delete disused DNA and to maintain megaplasmids. *Antonie Van Leeuwenhoek* **63**:39–43.
- HOWARD OCHMAN, reviewing editor
- Accepted August 30, 1999