Evolutionary Study of Amino Acid Substitution Patterns Associated with Accelerated Evolution in Endosymbiotic Bacteria

Jun-ichi Takeda[1], Takeshi Itoh[1], Tadashi Imanishi[1], Takashi Gojobori[1]

Keywords: endocellular symbiont, evolutionary rate, amino acid substitution, genome evolution

1 Introduction.

It was previously reported that the evolutionary rate of *Buchnera* spp., endocellular symbionts of aphids is about two times faster than that of *Escherichia coli*, their free-living cousin, and it was suggested that a major factor of this evolutionary rate acceleration is the enhanced mutation rate rather than fixation of slightly deleterious mutations [1]. On the other hand, other studies about the amino acid replacement indicates that the higher evolutionary rate in *Buchnera* is due to the slightly deleterious or beneficial mutation [2, 3]. If the neutral evolution of the enhanced mutation rate is a major factor, the amino acid substitution patterns in *Buchnera* may not change very much compared with those in *E. coli*, while accumulation of slightly deleterious or beneficial mutations would result in a drastic change of the substitution patterns (e.g. increase of radical substitutions). Here we examined the substitution patterns by using the Grantham matrix, which can be used to estimate differences of the amino acid substitution properties [4].

2 Materials and Methods.

Date set
We selected *Buchnera aphidicola* as an endocellular symbiont and *Escherichia coli* as a free-living bacterium. We used seven complete sequences of the following prokaryotes: *Buchnera aphidicola* str. APS (*Acyrhosiphon pisum*), *Buchnera aphidicola* (*Schizaphis graminum*), *Buchnera aphidicola* (*Baizongia pistaciae*), *Escherichia coli* K12 MG1655, *Escherichia coli* O157:H7 RIMD 0509952, *Haemophilus influenzae* Rd KW20, and *Salmonella enterica* subsp. *enterica serovar* Typhi CT18.

Comparison of amino acid substitutions
We used 85 genes, of which orthology was confirmed manually [1]. We made amino acid alignments of orthologs among *Buchnera aphidicola* str. APS, *Escherichia coli* K12 and *Haemophilus influenzae* by using ClustalW. Amino acid substitutions in each lineage were calculated by using the maximum parsimony method. Then, we examined the difference of amino acid property during the substitution by using the Grantham matrix, classifying them into four categories: very radical, radical, moderate, and conservative [6]. Likewise, we carried out the same analysis for *Buchnera aphidicola* str. APS, *Buchnera aphidicola* (*Schizaphis graminum*) and *Buchnera aphidicola* (*Baizongia pistaciae*) (outgroup), and for *Escherichia coli* K12, *Escherichia coli* O157 and *Salmonella enterica* (outgroup).

Motif search

[1] Biological Information Research Center, National Institute of Advanced Industrial Science and Technology, Time24 Bldg. 10F 2-45 Aomi, Koto-ku, Tokyo 135-0064, Japan. E-mail: jtakeda@jbirc.aist.go.jp

We searched for functional motifs in *Escherichia coli* K12 by using InterProScan, and examined the positions of the amino acid substitutions in terms of the locations of the motifs. The motif locations in *Buchnera aphidicola* str. APS were determined by using the motifs found in *Escherichia coli* K12.

Estimation of d_N/d_S

We estimated the nonsynonymous-synonymous ratios (d_N/d_S) of 85 genes between *Buchnera* spp. and between *Escherichia coli* and *Salmonella enterica* by using the Nei-Gojobori method.

3 Results.

If two-fold accelerated evolution in *Buchnera* is because of fixation of slightly deleterious or beneficial mutations, it was expected that the radical and very radical amino acid substitutions increase. However, the amino acid substitution patterns in *Buchnera* were not largely different from those in *Escherichia coli*. In addition, the amino acid substitute patterns were not very different between the inside and outside of functional motifs in the two species. Although d_N/d_S of *Buchnera* was larger than that of *Escherichia coli*, d_N/d_S of *Buchnera* is not different from other species except *Escherichia coli* [5]. These observations implies that the acceleration of the evolutionary rate is due to the enhanced mutation rate rather than the slightly deleterious or beneficial mutations.

References

[1] Itoh, T., Martin, W. and Nei, M. 2002. Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. *Proceedings of the National Academy of Sciences USA* 99:12944-12948.

[2] 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 bacterial symbiont. *Journal of Molecular Evolution* 53:377-386.

[3] Moran, A. N. 1996. Accelerated evolution and Muller's racher in endosymbiotic bacteria. *Proceedings of the National Academy of Sciences USA* 93:2873-2878.

[4] Grantham, R. 1973. Amino acid differences formula to help explain protein evolution. *Science* 185:862-864.

[5] Fares, M.-A., Barrio, E., Sabater-Munoz, B. and Moya, A. 2002. The evolution of the heat-shock protein GroEL from *Buchnera*, the primary endosymbiont of aphids, is governed by positive selection. *Molecular Biology and Evolution* 19:1162-1170.

[6] Li, W.-H., Wu, C.-I., and Luo, C.-C. 1985. A new method for Estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Molecular Biology and Evolution* 2:150-174.