Ureaplasma Urease Genes have Undergone a Unique Evolutionary Process

Hiromi Nishida*

Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

Abstract: *Ureaplasma*, a member of mycoplasmas, has a unique ATP synthesis system, which is coupled to the urea hydrolysis. Urease catalyzes the hydrolysis of urea into carbon dioxide and ammonia. Phylogenetic analyses of the urease genes indicated that *Ureaplasma* urease genes were not gained by recent horizontal transfer and have a unique evolutionary process. *Ureaplasma* unique ATP synthesis system leaded to breakdown of the glycolysis pathway. Some glycolytic genes are absent and some glycolytic genes are evolving under relaxed selection in *Ureaplasma*. Probably glycolytic genes can be used as an indicator of ATP synthesis system. Thus, the organisms that have incomplete glycolysis pathway or glycolytic genes evolving under relaxed selection would have an ATP synthesis system independently of the glycolysis.

Mycoplasmas are widespread in nature as parasites of mammals, reptiles, fishes, arthropods, and plants [1]. During the mycoplasma evolution, gene loss has occurred frequently, resulting in very small genome size [1-3]. The reductive evolution of mycoplasmas is still in progress. The genus *Ureaplasma* is a member of mycoplasmas, which generates 95% of its ATP using the hydrolysis of urea [4]. Growth of *Ureaplasma* is dependent on urea [5]. This unique ATP synthesis is not found in the other mycoplasmas. In fact, key enzymes in the glycolytic pathway are absent in *Ureaplasma* [6]. In addition, some glycolytic genes of *Ureaplasma* are evolving under relaxed selection [7, 8]. Thus, the glycolysis pathway is collapsing in *Ureaplasma*.

is coupled to ATP synthesis in *Ureaplasma* [9]. This unique system of *Ureaplasma* leaded to breakdown of the glycolysis pathway. For example, *Ureaplasma* does not have any genes encoding glucose-6-phosphate isomerae [10]. It suggests that the glycolysis system had been important for ATP synthesis rather than glucose metabolism at least in *Ureaplasma* (Fig. **1**). Generally the TCA (tricarboxylic acid) cycle is linked to the glycolysis pathway, which has played an important role in ATP synthesis of many organisms and conserved in the course of evolution. On the other hand, the urea hydrolysis is not linked to the glycolysis pathway. *Ureaplasma* generates ATP through the urea hydrolysis not through the glycolysis. Probably *Ureaplasma* glycolysis pathway was not able to be

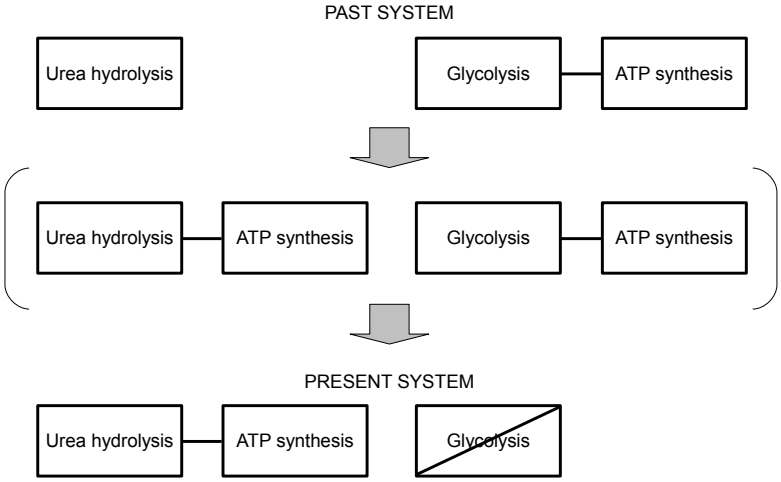


Fig. (1). Model of change of ATP synthesis system in *Ureaplasma*. In *Ureaplasma*, the ATP synthesis through the glycolysis had been changed to that through the urea hydrolysis. The glycolysis pathway was broken after the urea hydrolysis was coupled to ATP synthesis.

Urea is hydrolyzed into carbon dioxide and ammonia in many organisms. However, it is unique that the urea hydrolysis broken before the urea hydrolysis was coupled to ATP synthesis. Therefore, after the coupling of the urea hydrolysis and ATP synthesis, dominant ATP synthesis had been changed from through the glycolysis to through the urea hydrolysis during the *Ureaplasma* evolution (Fig. **1**).

*Address correspondence to this author at the Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan;
E-mail: hnishida@iu.a.u-tokyo.ac.jp

1876-3928/09

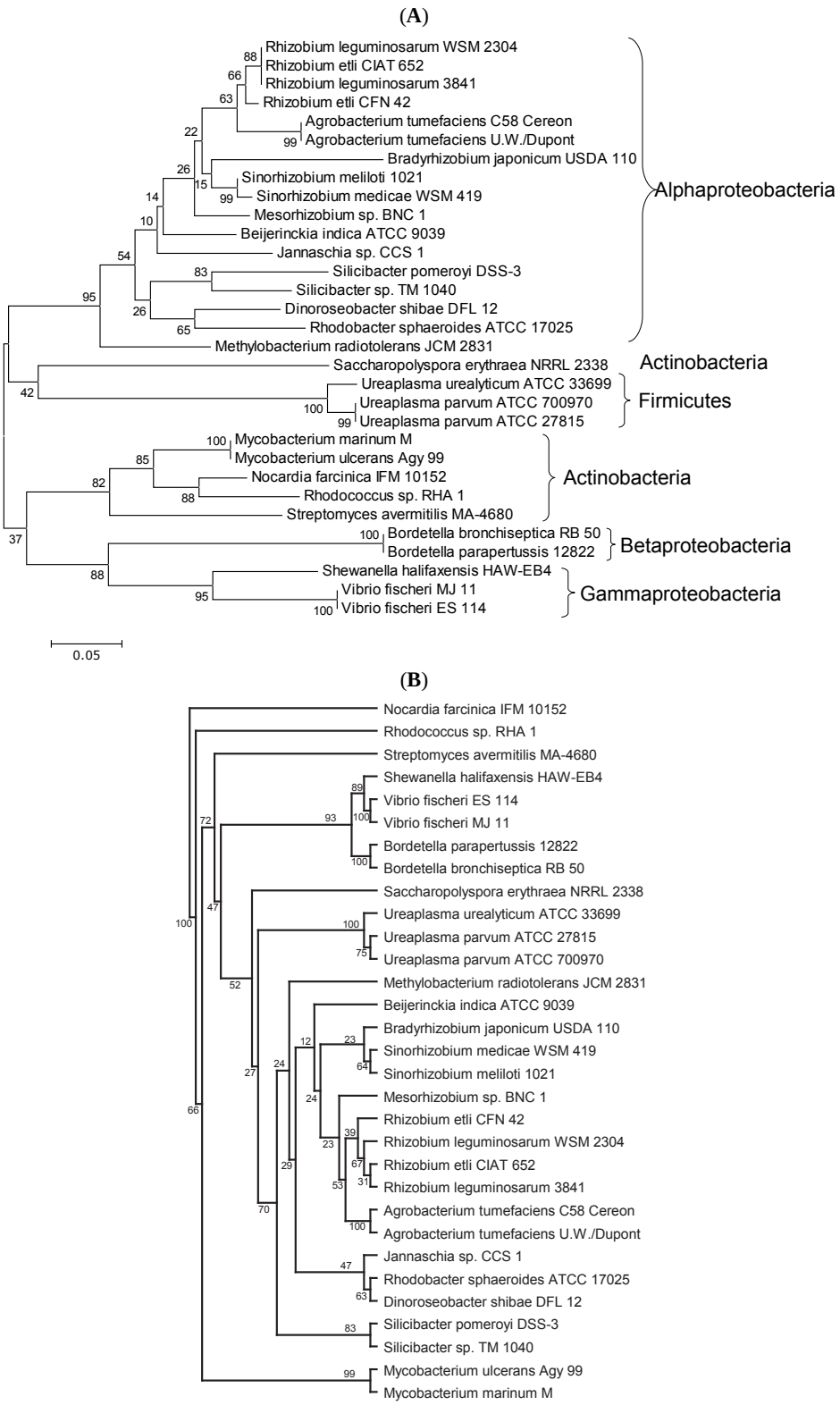
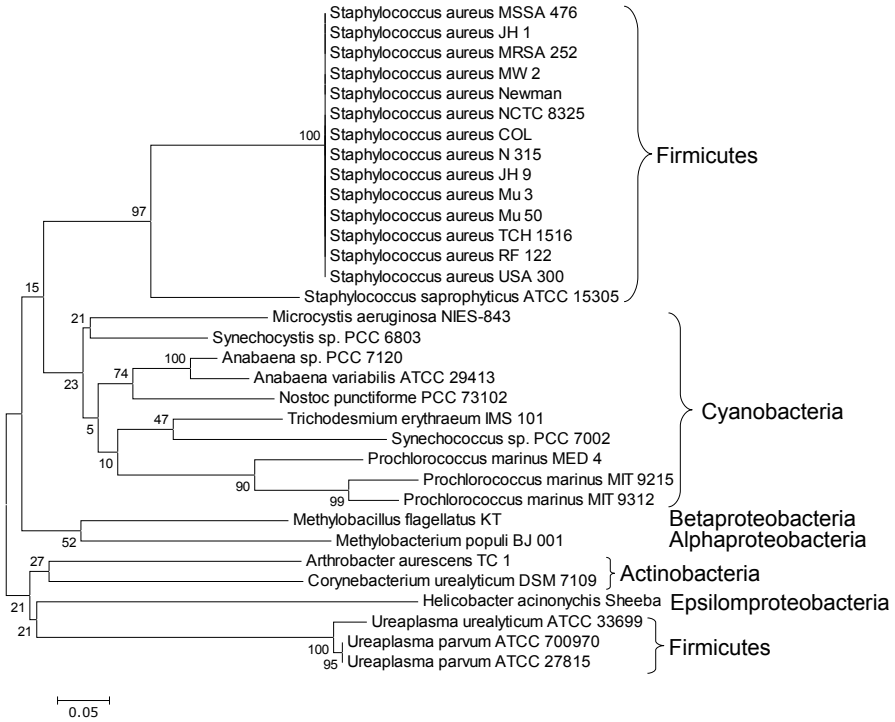


Fig. (2). Phylogenetic relationships among *Ureaplasma* UreA and the homologues. (**A**) Neighbo r-joining tree. The BLAST program was used to search the GenomeNet website for proteins homologous to *Ureaplasma* UreA with *E*-value < 10^{-23}. *Ureaplasma* UreA and 30 homologous proteins were multiple aligned using the CLUSTAL W program. Phylogenetic tree was reconstructed, based on the multiple-alignment with complete deletion of gap sites using the neighbor-joining method of MEGA software with 1000 bootstrap replicates. (**B**) Maximum likelihood tree. Phylogenetic tree was reconstructed using the maximum likelihood method of the PHYLIP program with 100 bootstrap replicates. The JTT model was used as the model of amino acid substitution. Number of times to jumble in the PROML program was 2.

(**A**)



(**B**)

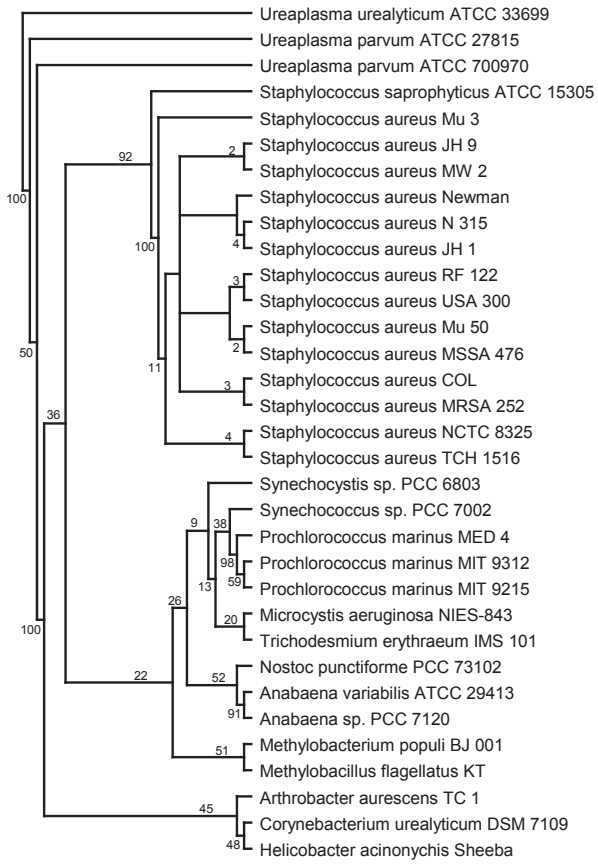


Fig. (3). Phylogenetic relationships among *Ureaplasma* UreB and the homologues. (**A**) Neighbor-joining tree. The BLAST program was used to search the GenomeNet website for proteins homologous to *Ureaplasma* UreB with *E*-value < 10^{-25}. *Ureaplasma* UreA and 32 homologous proteins were multiple aligned using the CLUSTAL W program. Phylogenetic tree was reconstructed, based on the multiple-alignment with complete deletion of gap sites using the neighbor-joining method of MEGA software with 1000 bootstrap replicates. (**B**) Maximum likelihood tree. Phylogenetic tree was reconstructed using the maximum likelihood method of the PHYLIP program with 100 bootstrap replicates. The JTT model was used as the model of amino acid substitution. Number of times to jumble in the PROML program was 2.

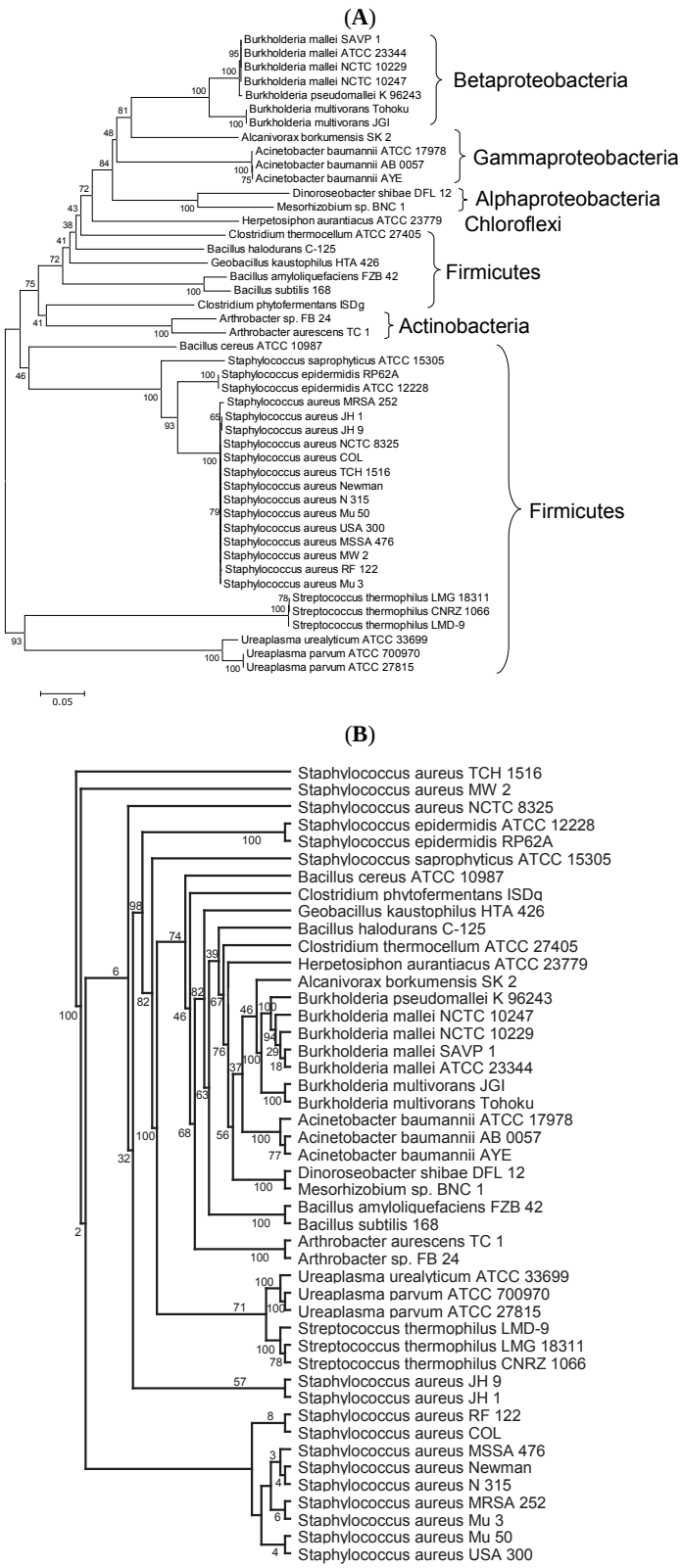


Fig. (4). Phylogenetic relationships among *Ureaplasma* UreC and the homologues. (**A**) Neighbor-joining tree. The BLAST program was used to search the GenomeNet website for proteins homologous to *Ureaplasma* UreC with *E*-value < 10^{-170}. *Ureaplasma* UreA and 45 homologous proteins were multiple aligned using the CLUSTAL W program. Phylogenetic tree was reconstructed, based on the multiple-alignment with complete deletion of gap sites using the neighbor-joining method of MEGA software with 1000 bootstrap replicates. (**B**) Maximum likelihood tree. Phylogenetic tree was reconstructed using the maximum likelihood method of the PHYLIP program with 100 bootstrap replicates. The JTT model was used as the model of amino acid substitution. Number of times to jumble in the PROML program was 2.

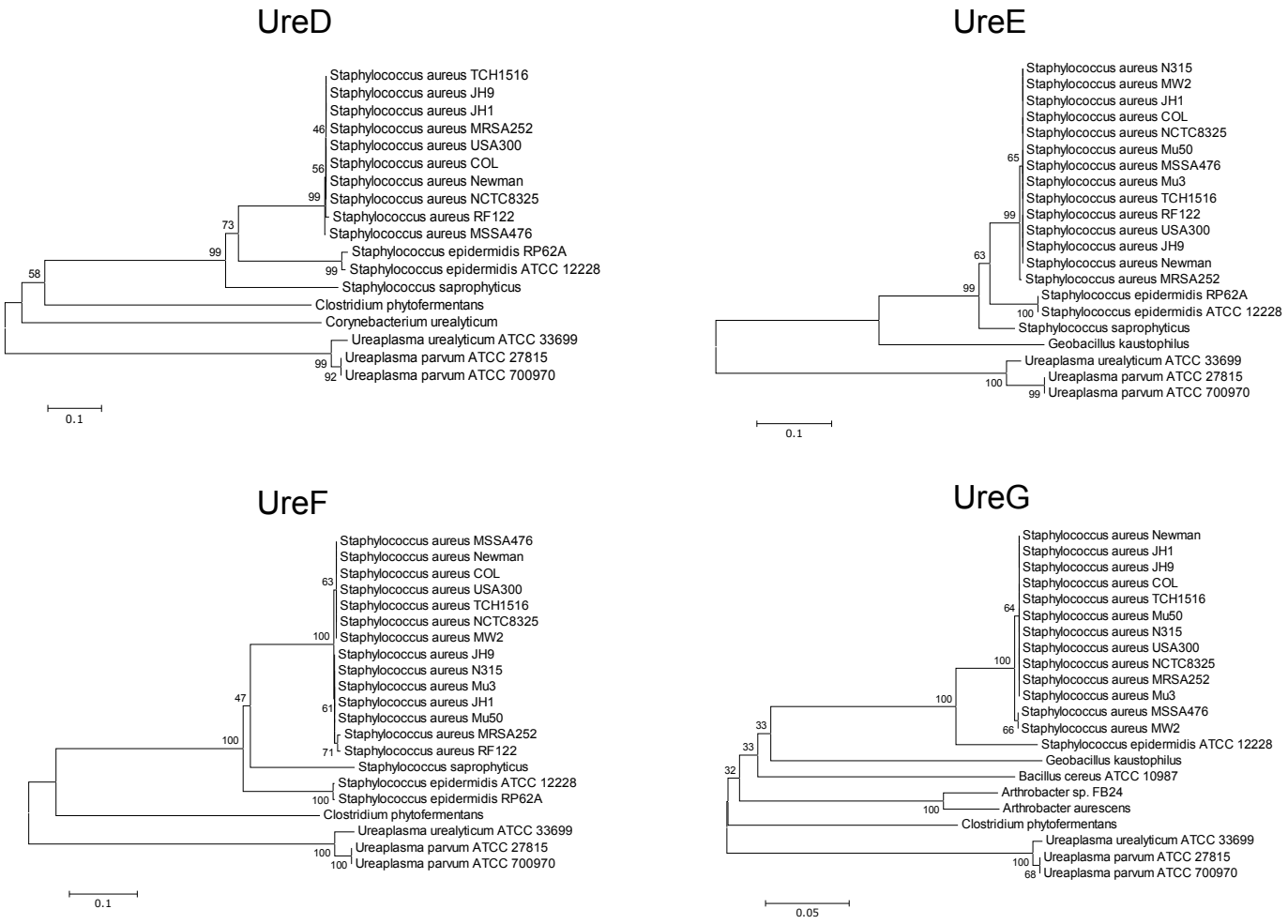


Fig. (5). Phylogenetic relationships among *Ureaplasma* UreD and the homologues, UreE and the homologues, UreF and the homologues, and UreG and the homologues. The BLAST program was used to search the GenomeNet website for proteins homologous to *Ureaplasma* UreD, UreE, UreF, and UreG of *Ureaplasma parvum* ATCC 27815 with *E*-value < 10^{-35}, 10^{-25}, 10^{-40}, and 10^{-77}, respectively. This search led to identification of 17 UreD homologues, 20 UreE homologues, 20 UreF homologues, and 21 UreG homologues. The 18 UreD, 21 UreE, 21 UreF, and 22 UreG proteins were multiple aligned using the CLUSTAL W program on the GenomeNet website. Phylogenetic tree was reconstructed, based on the multiple-alignment with complete deletion of gap sites using the neighbor-joining method of MEGA software with 1000 bootstrap replicates.

Urease catalyzes the hydrolysis of urea into carbon dioxide and ammonia. Urease has three core subunits, UreA, UreB, and UreC. Generally these three core subunits and the related accessory proteins UreD, UreE, UreF, and UreG are clustered on the genome. The order of the *Ureaplasma* (belongs to the Firmicutes) urease genes in the cluster is identical to that of the *Bacillus* sp. TB-90 (belongs to the Firmicutes) urease genes [11], suggesting that an ancestor of the Firmicutes had the urease gene cluster. Some species had lost some urease genes during the evolution. For example, *Bacillus subtilis* contains urease structural genes but lacks the accessory genes typically required for the maturation of urease [12]. On the other hand, the mycoplasmas except for *Ureaplasma* completely lack urease genes. I have some questions. Was the *Ureaplasma* urease genes (gene cluster) transferred horizontally? Or did the mycoplasmas except for *Ureaplasma* lose urease genes in the course of the mycoplasma evolution? Did the *Ureaplasma* urease have a unique evolutionary process? In order to answer these questions, I compared the amino acid sequences of urease genes of *Ureaplasma* with the homologous proteins in this study.

The BLAST program was used to search the GenomeNet website (http://www.genome.jp) for proteins homologous to UreA, UreB, UreC, UreD, UreE, UreF, and UreG of *Ureaplasma parvum* ATCC 27815 with *E*-value < 10^{-23}, 10^{-25}, 10^{-170}, 10^{-35}, 10^{-25}, 10^{-40}, and 10^{-77}, respectively. This search led to identification of 30 UreA homologues, 32 UreB homologues, 45 UreC homologues, 17 UreD homologues, 20 UreE homologues, 20 UreF homologues, and 21 UreG homologues. The 31 UreA, 33 UreB, 46 UreC, 18 UreD, 21 UreE, 21 UreF, and 22 UreG proteins were multiple aligned using the CLUSTAL W program on the GenomeNet website. Phylogenetic trees were reconstructed, based on the multiple-alignments with complete deletion of gap sites using the neighbor-joining method of MEGA software [13] with 1000 bootstrap replicates. In addition, phylogenetic trees of UreA, UreB, and UreC proteins were reconstructed, based on the multiple alignments with complete deletion of

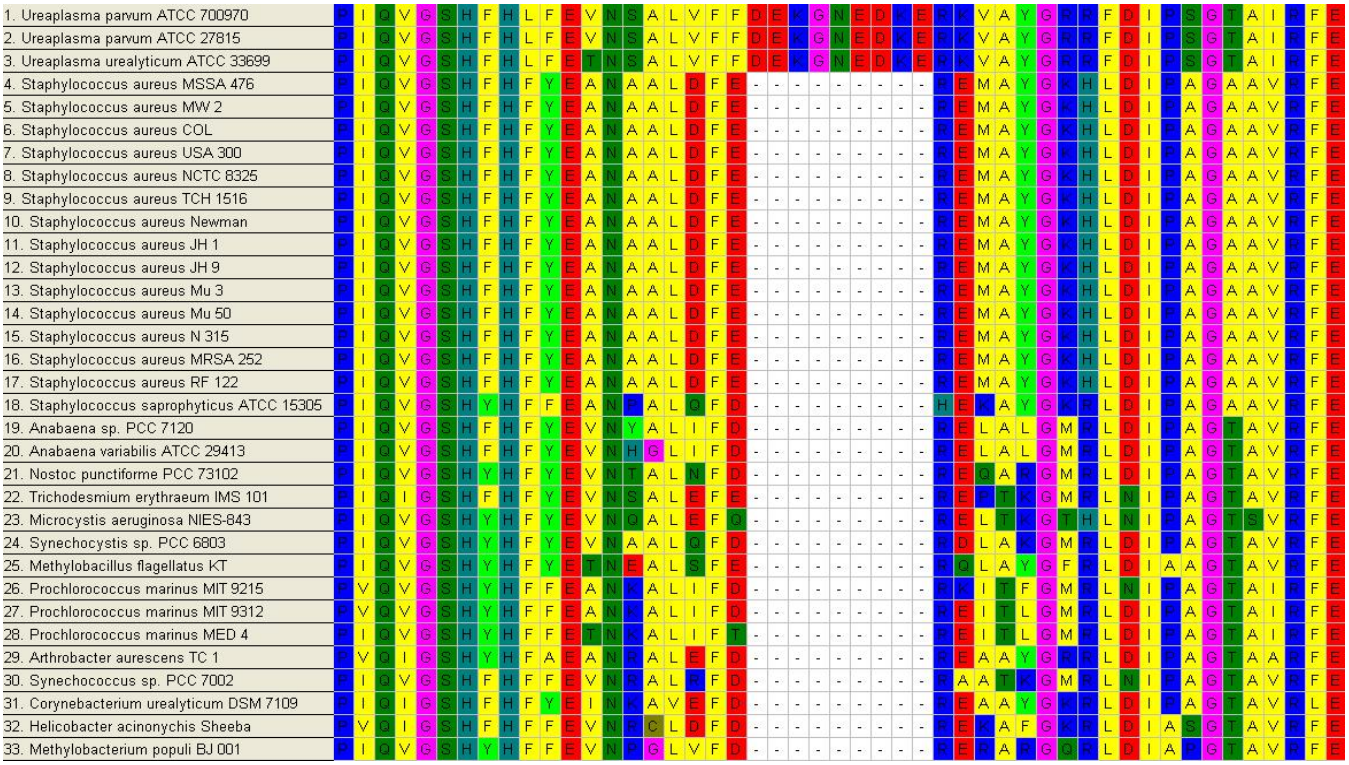


Fig. (6). *Ureaplasma* UreB specific amino acids insertion region. *Ureaplasma* has a specific 9-amino acids insertion (positions 65-73) in the multiple alignment of 33 UreB proteins.

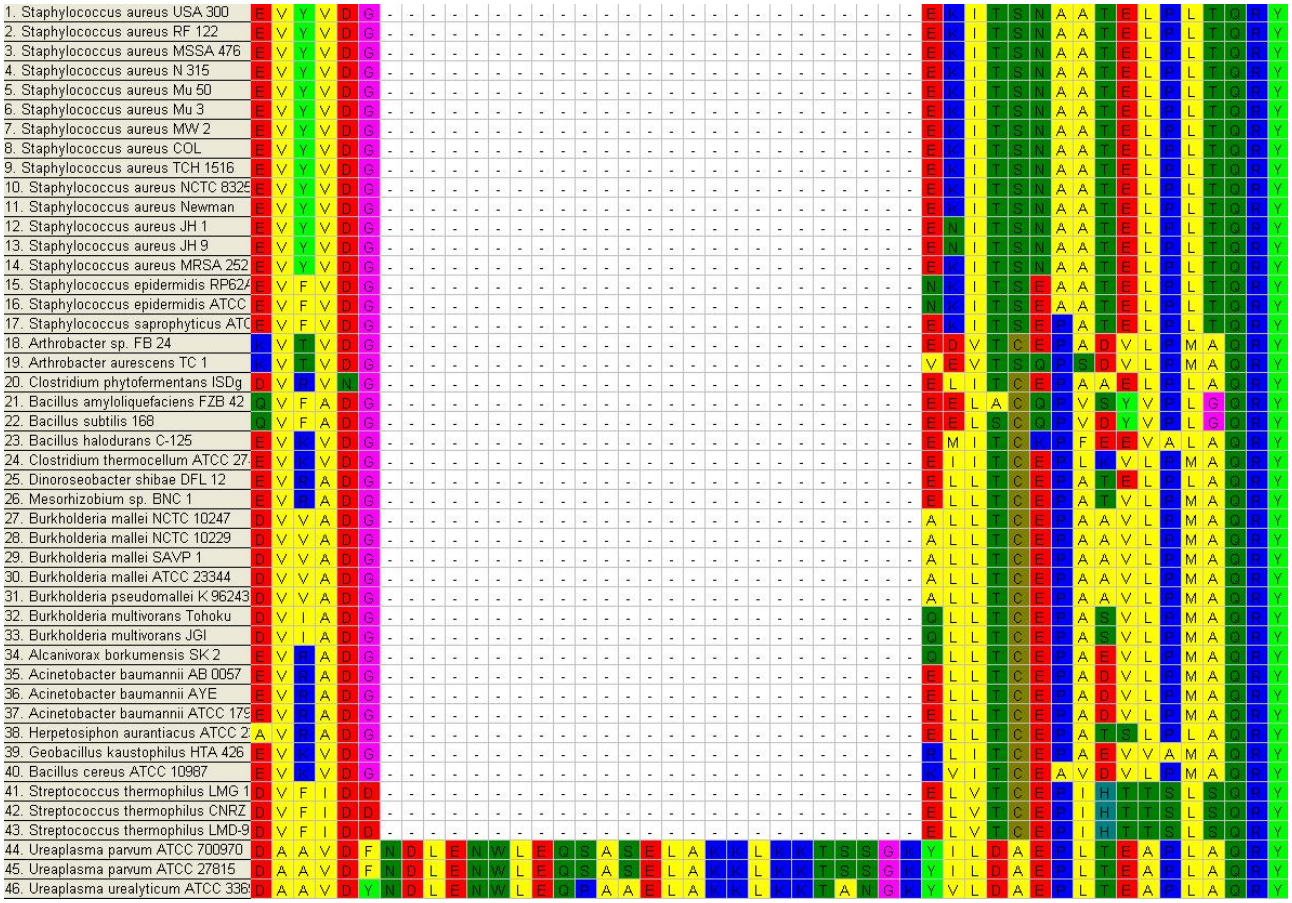


Fig. (7). *Ureaplasma* UreC specific amino acids insertion region. *Ureaplasma* has a specific 25-amino acids insertion (positions 554-578) in the multiple alignment of 33 UreC proteins.

gap sites using the maximum likelihood method of the PHYLIP program (http://evolution.genetics.washington.edu/phylip.html) with 100 bootstrap replicates. The JTT model was used as the model of amino acid substitution. Number of times to jumble in the PROML program was 2.

The phylogenetic analyses showed that the KEGG database [14] used in this study did not have closely related protein to UreA and UreB of *Ureaplasma* (Figs. **2A**, **B** and **3A**, **B**). On the other hand, *Ureaplasma* UreC is closely related to *Streptococcus* UreC with 93% bootstrap support in the neighbor-joining tree (Fig. **4A**) and 71% support in the maximum likelihood tree (Fig. **4B**). The diverging points of UreA, UreB, and UreC of *Ureaplasma* are very deep in the neighbor-joining trees (Figs. **2A**, **3A**, **4A**), strongly suggesting that *Ureaplasma* urease genes (gene cluster) were not gained by recent horizontal transfer. It was also supported by the phylogenetic trees of urease accessory proteins (Fig. **5**). Interestingly, *Ureaplasma* UreB and UreC have *Ureaplasma*-specific amino acids insertions (Figs. **6**, **7**). Thus, *Ureaplasma* urease gene cluster has a unique evolutionary process and *Ureaplasma* has a unique urea hydrolysis system coupling to ATP synthesis. Maybe *Ureaplasma* ancestor had used urease in order to live in the urea-rich environment. During the evolution, the urea hydrolysis was coupled to ATP synthesis system. After the event, glycolysis pathway has not been essential for ATP synthesis in *Ureaplasma*. Some glycolytic genes are absent and some are evolving under relaxed selection in *Ureaplasma*. Probaly the establishment of the unique ATP synthesis system triggered those changes. If so, glycolytic genes can be used as an indicator of ATP synthesis system. Thus, the organisms that have incomplete glycolysis pathway or glycolytic genes evolving under relaxed selection would have a unique ATP synthesis system independently of the glycolysis.

REFERENCES

[1] Razin S, Yogev D, Naot Y. Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev 1998; 62: 1094-156.

[2] Jaffe JD, Stange-Thomann N, Smith C, *et al*. The complete genome and proteome of Mycoplasma mobile. Genome Res 2004; 14: 1447-61.

[3] Oshima K, Kakizawa S, Nishigawa H, *et al*. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat Genet 2004; 36: 27-9.

[4] Smith DGE, Russell WC, Ingledew WJ, Thirkell D. Hydrolysis of urea by Ureaplasma urealyticum generates a transmembrane potential with resultant ATP synthesis. J Bacteriol 1993; 175: 3253-8.

[5] Kenny GE, Cartwright FD. Effect of urea concentration on growth of Ureaplasma urealyticum (T-strain mycoplasma). J Bacteriol 1977; 132: 144-50.

[6] Pollack JD. Ureaplasma urealyticum: an opportunity for combinatorial genomics. Trend Microbiol 2001; 9: 169-75.

[7] Oshima K, Nishida H. Phylogenetic relationships among mycoplasmas based on the whole genomic information. J Mol Evol 2007; 65: 249-58.

[8] Oshima K, Nishida H. Detection of the genes evolving under Ureaplasma-specific selection. J Mol Evol 2008; 66: 529-32.

[9] Romano N, La Licata R, Russo Alesi D. Energy production in Ureaplasma urealyticum. Pediatr Infect Dis 1986; 5: S308-12.

[10] Glass JI, Lefkowitz EJ, Glass JS. The complete sequence of the mucosal pathogen. Ureaplasma urealyticum. Nature 2000; 407: 757-62.

[11] Neyrolles O, Ferris S, Behbahani N, Montagnier L, Blanchard A. Organization of Ureaplasma urealyticum urease gene cluster and expression in a suppressor strain of *Escherichia coli*. J Bacteriol 1996; 178: 647-55.

[12] Kim JK, Mulrooney SB, Hausinger RP. Biosynthesis of active Bacillus subtilis urease in the absence of known urease accessory proteins. J Bacteriol 2005; 187: 7150-54.

[13] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007; 24: 1596-99.

[14] Kanehisa M, Araki M, Goto S, *et al*. KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008; 36: D480-D484.

Received: December 1, 2008 Revised: December 16, 2008 Accepted: December 17, 2008