Evaluation of Pathogenic Bacterial Taxonomy after Complete Genome Analysis

Takayuki Ezaki
Department of Microbiology, Gifu University Graduate school of Medicine
Yanagido 1-1, Gifu, Japan

Most of high rick pathogens were described in early period of 20th century and their taxonomy were exposed to many different level of classification tools. After the complete genome information of these high risk pathogens and related species, their classification was revealed to carry many taxonomic problems. Most environmental bacterial species shares more than 98,7% of 16S rDNA sequence similarities, however, the criteria cannot differentiate high rick pathogens from closely related species. *Bacillus anthracis, B.cereus, B.thuringens*is and *B. mycoides* shares more than 99.9% sequence similarities. *Escherichia coli* and 4 species of genus *Shigella* shares almost 100% of 16S rDNA sequence. House keeping genes associated to their replication, modification and protein synthesis were often used to differentiate closely related species because of their SNP variation were bigger than 16S rDNA sequence.

ropB, gyrB, and *dnaJ* sequences are often used to differentiate closely related species but SNP range within a species ranges from 0 to 5%. To evaluate species range of these genes were used to be difficult because most sequences were partial.

After the accumulation of whole genome sequences, house keeping genes of many strains in a species became available. We extensively analyzed protein sequence variation among strains of the high risk pathogens and the relatives. We focused on the two group of pathogens. *Eschericha coli* group and *Bacillus cereus-anthracis* groups. Among 40 strains of *Eschericha coli* genomes, ranges of house keeping gene sequence and protein sequence were calculated. Using these data, taxonomic position of *Escherichia coli, E. fergsonii, E. albertii* , *E. vulneris*, *E. hermanii* and four species of *Shigella* are discussed.

Bacillus cereus, Bacillus anthracis, Bacillus thuringensis, Bacillus mycoides, B. pseudomycoides, and B.weihenstephanensis pause another taxonomic problems. These share almost identical 16S rDNA sequences. Whether we shall accept these species as independent species, or reclassify them as pathovars in a single species will be discussed after their complete genome comparison at sequence or protein level.