S3-3

Microbial Genome Annotation System

JUNG Jongsun
Nanormics, Inc.

The annotation process of a full genome requires a whole set of bioinformatic tools as well as accumulated and established biological data on sequences and structures of the genes and the proteins. With the tools and the databases, in silico genome annotation swiftly assigns value-added meaning to sequence data that would otherwise be treated manually by a human expert for a long period of time. Here we introduce a microbial genome annotation system that is equipped with such tools and databases. The annotation system consists of three topics; structural genomics, functional genomics and comparative genomics. Given a full genome sequence, protein coding regions are identified and their putative functions for those regions are, then, defined based on a series of conventional anannotation procedures and then, predicting uncharacterized operons and building a global regulatory network are performed. The network for a given genome includes clusters for primary and secondary metabolites, regulon, modulon and stimulon as well as protein pairs for protein-protein interaction. So far, we have re-annotated three full genomes, E.coli-k12, S.coelicolor-A3 and S.avermitilis, and annotated one full genome, S.peucetius. Comparison between microbial genomes, biological consequence and analysis of the global regulatory network will be discussed.

S3-4

Complete Genome Sequence of *Vibrio vulnificus* CMCP6, the Etiological Agent of a Shellfish-Associated Fatal Septicemia

RHEE Joon Haeng[1,2], JEONG Hae Young[3], KIM Soo Young[1,2], LEE Shee Eun[1,2], MOON Young Ho[3], KIM Jae Jong[3], CHUNG Sun Sik[1], and CHOY Hyon E.[1]
[1]Genome Research Center for Enteropathogenic Bacteria and Department of Microbiology,
[2]National Research Laboratory of Molecular Microbial Pathogenesis, Chonnam National University Medical School
[3]Genotech Corp., Daejeon, Korea

Vibrio vulnificus is a halophilic gram-negative estuarine bacterium that causes a fatal septicemia and necrotizing wound infections. The whole genome sequence of a Korean clinical isolate CMCP6 was determined by using the shotgun sequencing strategy. Two types of sh otgun library (2-3 and 4-5 kbp) and a fosmid genome library were employed for the shotgun sequencing and finishing. The genome consisted of two circular chromosomes of 3,281,945 bp and 1,844,853 bp, 20% bigger than that of *V. cholerae* and similar size with that of *V. parahaemolyticus,* that together encode 4,894 open reading frames (ORFs). Among the predicted ORFs, 3,791 (77.5%) could be assigned into COGs (cluster orthologous groups) reported in the GenBank database, and 1,955 (39.9%) could be assigned with no previously known functions. The genes associated with essential cell functions (such as translation and ribosome biogenesis, cell division, secretion, nucleotide transport and metabolism, and coenzyme metabolism) had predilection to the larger chromos ome. In contrast, the small chromosome showed bias to the genes for carbohydrate transport and metabolism, defense mechanisms, transcription, signal transduction, secondary metabolism, and inorganic ion transport and metabolism. Most of the exotoxin genes r eported to be associated with virulence of *V. vulnificus* were located on the small chromosome. The super-integron island, supposed to function as a gene capture system and reported on the *V. cholerae* small chromosome, was discovered on the larger chromosome. The super-integron island showed lower G+C ratio, spanned 152 kbp, and contained 131 ORFs encoding mostly hypothetical proteins. The *V. vulnificus* genome sequence provides a scaffolding for understanding how the pathogen survives in the environment and causes fatal septicemia.