

# 2<sup>nd</sup> International Congress on **Bacteriology & Infectious Diseases**

November 17-19, 2014 DoubleTree by Hilton Hotel Chicago-North Shore, USA



***Michael Y Galperin***

National Center for Biotechnology Information, USA

## **Functional annotation of bacterial genomes, one COG at a time**

The availability of complete genome sequences allows us to take an unbiased look at the live cell and evaluate the degree of our understanding of the key cellular processes. Functional annotation of the recently completed genomes showed that only a small fraction of proteins encoded in any given genome has been ever studied experimentally or could be studied in detail any time soon. Therefore, computational analysis remains the key approach to functional assignments of the majority of genes in any prokaryotic or eukaryotic cell. For such approach to work, it is important to understand the limits for functional predictions based on sequence similarity and/or membership in protein (super) families. This talk will discuss the current status of functional annotation for various bacteria, from well-characterized model organisms to uncultured parasites and symbionts, where the genome sequence provides the only clue to the organism's physiology. Using the recently updated COG names, the recent progress in characterizing the functions of "conserved hypothetical" proteins and the challenges in completing the "parts list" in the best-studied genomes will be discussed. It now appears that the number of truly enigmatic proteins is relatively small, particularly in the reduced genomes of bacterial pathogens. In contrast, the number of uncharacterized genes in free-living bacteria and most eukaryotes remains quite large and their functions remain obscure. Hypothesis is that many of the remaining uncharacterized genes have "house-cleaning" (as opposed to house-keeping) function, hydrolyzing or detoxifying the side products of bacterial metabolism.

## **Biography**

Michael Y Galperin has received his PhD at the Lomonosov Moscow State University in Russia and Postdoctoral training at the University of Louisville and University of Connecticut. He has been at the NCBI Computational Biology Branch since 1996, first as a GenBank Fellow, then Staff Scientist, and currently as a Lead Scientist. He has published more than 170 research papers, reviews and book chapters and is an author of a textbook on comparative genomics. He serves as the editor of the *Nucleic Acids Research* annual Database Issue and Editor of the Genomics Updates section in *Environmental Microbiology*.

[galperin@ncbi.nlm.nih.gov](mailto:galperin@ncbi.nlm.nih.gov)