Identifying the micro from the peta: tales of big data in the micro world

Nikos Kyrpides,
Prokaryotic Super Program
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek CA 94598 USA

Shotgun sequencing of environmental samples has revealed a new universe of microbial communities (metagenomes) involving previously uncultured organisms, which is expected to further advance our understanding the structure and function of entire microbial communities and expand our current knowledge of genetic and functional diversity of individual micro-organisms. Although most metagenomic studies aim to unveil the genetic and functional novelty present in various environments, current best practices in metagenome analysis tools, and workflows, are suffering from a number of pitfalls, impairing the fulfillment of this objective. I will present some of our approaches enabling the exploration of large metagenomic datasets and the discovery of novelty. Specifically I will present approaches and results for the discovery of novel phylogenetic lineages as well as the exploration of the viral dark matter.