Journal of Molecular and Genetic Medicine

Kalia, J Mol Genet Med 2015, 9:2
DOI: 10.4172/1747-0862.1000E105

Editorial Open Access

Let's Explore the Latent Features of Genes to Identify Bacteria

Vipin Chandra Kalia*

Microbial Biotechnology and Genomics, CSIR-Institute of Genomics and Integrative Biology, India

***Corresponding author:** Vipin Chandra Kalia, Microbial Biotechnology and Genomics, CSIR-Institute of Genomics and Integrative Biology, Delhi University Campus, Mall Road, Delhi - 110007, India, Tel: 091- 11- 27666156; 27666157; Fax: 091- 11 - 27667471; 27416489; E-mail: vckalia@igib.res.in, vc_kalia@yahoo.co.in

Received date: April 13, 2015; **Accepted date:** April 15, 2015; **Published date:** April 22, 2015

Editorial

Bacterial world is immensely great and interesting. In spite of the fact that microbes existed much before their presence was demonstrated. It was because of the vision and abilities of the great men like Prof. Robert Hooke and Prof. Antonie van Leeuwenhoek, that we have learnt all we know today about microbial world. Their discoveries laid the basis for the Microbiology of today. The diversity of their features has been the focus of research of microbiologists. The orbit of the magnitude of the diversity of their metabolic, phenotypic, genomic characteristics allows them to survive in a wide diversity of climatic conditions. In view of their diverse biotechnological applications in virtually all spheres of human activities, the curiosity has driven us to attend to their genuine individuality. Beginning with morphological features as the basis of separating them from one another, there has been a shift to look deeper into their biochemical and genetic characteristics. The advent of molecular biology techniques has completely overturned the concept of bacterial taxonomy and their evolutionary story. In the innovative era of Molecular biology and Bioinformatics, the vision of Prof. Carl R. Woese, took the tenets of evolutionary biology and taxonomy to a new peak. The Microbiologists of today have taken advantage of the molecular techniques developed over a few decades towards the end of the previous century and have entered into the era of genomics and metagenomics.

The troubled waters of the microbial world

Almost all bacterial genera are composed of members which vary from each other at least in a few characteristics. On the evolutionary scale, it is relatively easy to distinguish organisms which are distantly placed. However, there is too much of heterogeneity in terms of their phenotype and genotypes that it becomes difficult to categorize them unambiguously. In this write up we will take up only a few cases, to illustrate the latent features of *rrs* and other metabolic genes, which can be easily exploited for identifying even closely related organisms within a genus: (i) *Bacillus*, (ii) *Clostridium*, and (iii) *Pseudomonas*.

Bacillus is one of those few organisms, whose applications can be seen in almost all spheres of life: biofuels, crop protection, biopharmaceuticals - peptide antibiotics. In the same flow it may be relevant to state that *Bacillus* as contaminants in industrial processes, food poisoning, and probiotics are the most troublesome. Here, it becomes imperative to identify them with utmost precision. *B. subtilis* group members, intraspecies diversity of *B. licheniformis*, *B. cereus* group members, *B. sphaericus* are, to cite a few cases, which have been the major trouble spots in the genus *Bacillus*. Whole genome sequencing of *B. halodurans* has revealed them to possess genes coding for the enzymes transposases and recombinases, showing high similarity to those present in *Anabeana*, *Rhodobacter*, *Lactococcus*, *Enterococcus*, *Clostridium*, *Staphylococcus* and *Yersinia* species. Comparative genome analysis of B. halodurans and B. subtilis provided evidences (genome size, genomic DNA G+C content and the physiological properties - ATPases, ABC transporter genes, etc. of their being very close to each other, in terms of. The need to reclassify certain organisms previously belonging to Bacillus has resulted in shifting them to new genera: *Aneurinibacillus, Brevibacillus*, *Paenibacillus*, *Ureibacillus* and *Virgibacillus*: (i) *B. thermoleovorans*, *B. sterothermophilus*, *B. thermoglucosidasius*, and *B. kaustophilus*, as *Geobacillus*, (ii) *B. psychrophilus*, *B. pasteurii* and *B. globisporus*, as *Sporosarcina*, (iii) whereas *B. marinus* as *Marinibacillus marinus* [1].

Clostridium is another, widely studied bacteria, whose biotechnological applications range from solvents, enzymes, biofuels to deadly toxin producing pathogens. *Clostridium* species are phenotypically and phylogenetically heterogeneous group comprising of spore and non-spore formers, gram-negative and gram-positive, toxin and non-toxin producers. It thus becomes quite tedious to identify them. The whole process of identification is complicated also by the fact that the GC content varies from as low as 24 mol% in *Clostridium perfrigens* to as high as 58 mol% in the case of *Clostridium barkeri*. In comparison to a narrow GC content in most bacterial genomes, this range is quite wide to justify the presence of such organisms in a single genus. Another characteristic which pose a big problem is the presence of multiple copies of *rrs* gene, leading to over estimation of species [2].

Pseudomonas is one of those genus, which has been well studied. It has also been realized that it comprises of genetically and metabolic diverse organisms, which can be used for degrading a wide range of organic substrates. Initially, a large number of organisms were categorized as *Pseudomonas*, so much so that the genus was labeled as 'dumping ground'. It is thus no surprise that many of the organisms have undergone taxonomic revisions. Segregation of *Pseudomonas* has been based on the phylogenetic relationship of 4 housekeeping genes: *rrs*, *gyrB*, *rpoB* and *rpoD*. It enabled formation of relatively distinct groups: (i) cluster I members were the following species of *Pseudomonas*: *P. aeruginosa* along with *P. flavescens*, *P. mendocina*, *P. resinovorans*, and (ii) cluster II had the following organisms: *P. fluorescens* along with *P. chlororaphis*, *P. syringae* and *P. putida*. Another analysis based on the three genes except *rpoB* lead to a different scenario, where *P. aeruginosa* and *P. stutzeri* were together which was distinguishable from the group of *P. syringae*, *P. fluorescens* and *P. putida*. It thus urged researchers to look in to the relationships of organisms clubbed under *Pseudomonas* [3].

The modern taxonomic classification of microbes is based largely on the gene, which is conserved throughout the prokayotic domain: the 16S rRNA (*rrs*). The microbial taxonomy was given a new look and the nucleotide sequence of this gene has been so widely adapted that it has become a reference point for almost all practical purposes. The

Ribosomal Database Project (RDP) (https://rdp.cme.msu.edu/), which was developed as a small depository of a few hundred *rrs* sequences around two decades ago has more than 3.0 million entries (RDP Release 11, Update 3::September 17, 2014:: has 3,019,928 16S rRNAs:: 102,901 Fungal 28S rRNAs entries). The rapidly increasing magnitude of this database, based on the contributions made by researchers world-wide, is a clear reflection on the influence of the findings of Prof. Carl R. Woese. At times, the *rrs* gene sequence is not able to differentiate very closely related taxa. In such a scenario, one needs to resort to gene sequences which code for features such as heat shock proteins, ATPase-β-subunit, RNA polymerases or recombinase, etc. In certain cases, additional genes have been identified, which can be used exclusively for distinguishing members within a genus: (i) *rpoB* for *Mycobacterium*; (ii) *gyrB* for *Acinetobacter*, *Mycobacterium*, *Pseudomonas*, and *Shewanella*, (iii) *gyrA* gene for *Bacillus subtilis*, etc. A few methods generally used for identifying bacterials strains are: Amplified fragment length polymorphism (AFLP), DNA-DNA re association, Microarray, PCR-ribotyping, multi-locus sequence analysis, multi locus sequence typing, randomly amplified polymorphic DNA, and Restriction enzyme digestion (RE).

The latent features of 16S rDNA

The RDP data base used as reference to identify the newly sequenced 16S rDNA is limited by the fact that it can be used identify only to the extent to what has been known and is available. In order to identify the gene sequence which is yet to be seen by the database, it is difficult to visualize how to generate the evolutionary map and place the unknown. Efforts to resolve the potential problems existing among the different species of *Bacillus*, *Clostridium* and *Pseudomonas* revealed the presence of certain latent features in their 16S rDNA gene. The first step involved the generation of molecular makers was to develop a Phylogenetic Framework, which was composed of sequences, which delineated one species from another i.e. those sequences which could be used to demarcate the phylogenetic limits of all the known sequences within a species. The second step was to identify motifs (signatures sequences, 30 to 50 nucleotide in length), which were unique to a particular species and completely absent from all other species. The third feature, which validates the true identity of the 16S rDNA was the identification of Restriction Endonuclease (RE) which gives a unique digestion pattern: fragment lengths (nucleotides) and the order of their occurrence. These efforts helped in identification of organisms which were identified initially (by the inventor) only up to genus level. This humble beginning in identifying the latent features of those organisms which have been already well identified will help in future to identify and place them on the phylogenetic tree. In fact, these tools have been used to a small extent in certain studies; however, a complete study has been undertaken successfully by others to identify clinically important members of the genus *Streptococcus* [4,5].

References

1. Porwal S, Lal S, Cheema S, Kalia VC (2009) Phylogeny in aid of the present and novel microbial lineages: diversity in *Bacillus*. PLoS One 4: e4438.
2. Kalia VC, Mukherjee T, Bhushan A, Joshi J, Shankar P, et al. (2011) Analysis of the unexplored features of *rrs* (16S rDNA) of the Genus *Clostridium*. BMC Genomics 12: 18.
3. Bhushan A, Joshi J, Shankar P, Kushwah J, Raju SC, et al. (2013) Development of Genomic Tools for the Identification of Certain Pseudomonas up to Species Level. Indian J Microbiol 53: 253-263.
4. Huma N, Shankar P, Kushwah J, Bhushan A, Joshi J, et al. (2011) Diversity and polymorphism in AHL-lactonase gene (aiiA) of *Bacillus*. J Microbiol Biotechnol 21: 1001-1011.
5. Lal D, Verma M, Lal R (2011) Exploring internal features of 16S rRNA gene for identification of clinically relevant species of the genus Streptococcus. Ann Clin Microbiol Antimicrob 10: 28.