Indian J. Plant Genet. Resour. 29(3): 402-403 (2016)
DOI 10.5958/0976-1926.2016.00084.X

Insights from the Genome and Genome Informatics of Agriculturally Important Microorganisms

Alok K Srivastava
ICAR-National Bureau of Agriculturally Important Microorganisms, Maunath Bhanjan–275103, Uttar Pradesh, India

Microorganisms have been evolving on earth for at least 3.8 billion years. However, unlike animal or plant diversity, complete map of microbial diversity in a particular ecosystem is not available. Over the past few years, high-throughput sequencing has been revolutionized by the flooding of several next-generation sequencing (NGS) platforms. High-throughput next generation sequencing of microbial genomes and transcriptomes has changed the face of present era's microbiology. With the adventures of new technologies and methodologies, now it became possible to sequence a large microbial genome/ transcriptome within few days. Microsatellites such as simple sequence repeat markers (SSRs) are dispersed throughout the genome/ transcriptome which can be located and characterised through using the tools of genome informatics. It became easier to compare the multiple genomes through genome informatics. Genome wise comparison reveals the similarity and differences among the different organisms. Moreover, the recombination causes frequent genome rearrangements, horizontal transfer introduces new sequences into bacterial chromosomes, and deletions remove segments of the genome. These technologies are far more powerful than any existing molecular tools in diversity mapping and genome comparative analysis. However, our ability to accurately analyze these data relies on well-characterised, foundational reference data of microbial genomes. These sequencing techniques provide unique opportunity for researchers to build fundamental knowledge on microbial communities in the environment and understanding their role in management and protection of a biological system.

The high-quality throughput genome sequencing could be used to collect the real information and to utilise these insights in improving the important metabolic pathways to harness their potential in agriculture. A quick search of Genome online database (GOLD) showed that about 56,000 bio-samples, 19,200 studies, 56,000 sequencing projects and 39,400 analysis projects have been performed for microorganisms from the extreme environments. However, genome and transcriptome sequencing of the agriculturally important microorganisms (AIMs) is highly overlooked throughout the world; especially in India. However, several microbial genomes including archeal genomes sequenced in the country have given insights which can be utilized to decipher the coding sequences of key enzymes of different metabolic pathways, and changes in their expression profiles for better production of useful compounds/metabolites. Efforts are ongoing to perform the depth analysis of microbial genome sequences, and consequently to predict the sequences and structures of important proteins. The nucleotide sequences for key enzymes of different pathways can be utilized for the prediction of corresponding protein sequences and their structures through molecular modeling and simulations.

ICAR-NBAIM has sequenced the genomes of nine important bacterial strains that were previously characterised as potential nitrogen fixer rhizobia, stress mitigating salt / temperature tolerant organisms. Genome architecture of micro-organisms is the key targets to know that how the antibiotic resistance is being transferred or how the lysogenic phage get inserted into the bacterial genome and add the unusual characteristics. The comparison of whole genome sequence of six species of *Mesorhizobium* have shown that maximum numbers of singletons were found in *M. ciceri* ca181 and least in *M. australicum* 2075. Whole genome phylogeny analysis by different methods revealed that *M.loti* MAFF303099 and *M.huakuii* 7653R are in monophyletic clade and has a close relationship *M. ciceri* ca181. *Nif* gene analysis inferred that these are conserved during evolution. *M. ciceri* ca181 has maximum unique genes among six bacteria. The comparative whole genome analysis of these six *Mesorhizobium* has led to the identification of several proteins which are specific to particular strain and their roles are specific. In another study on the genome

---

**Author for Correspondence: Email- aloksrivastava@gmail.com* of a salt tolerant bacterium *Exiguobacterium profundum* PHM11, the presence of lysogenic phage DNA at five different places in the chromosome was recorded and recognized the portions of phage DNA which encode for the key proteins involved in the transport across the plasma membrane. We identified a *virD4* gene which is an ATPase that helps in transport across the membrane. Not only genome, but also the transcriptome analysis of microbial genomes are emerging area, and somehow related to the genome informatics. Researchers are also trying to elucidate that how the salt or drought stress can affect the plant physiology or how these stressors induced effects can be mitigated through microbial inoculations. Further, NGS driven amplicon based metagenomics have revolutionized the field of microbiology by providing a culture independent technique to identify and assess microbial diversity. The ability to sample environments for complex fungal metagenomes is rapidly becoming a reality and will play an important part in efficient utilisation of fungi for agricultural, industrial and climate management purposes. However, the metagenomic approach still has inherent limitations as it cannot distinguish live from dead or active from inactive microbial cells.

Genome and genome informatics, therefore, is an emerging area in the agricultural science. It could be only a possible way to get the real inferences which are present in the genomes of AIMs which can be utilized to decipher the coding sequences of key enzymes of different metabolic pathways, and changes in their expression profiles when coming in the contact with agricultural crops. Different aspects of agricultural biotechnology have added their inputs; however, plant/microbial genomics and genome informatics has contributed significantly. Besides these contributions, still numbers of questions have remained to be answered such as (a) can *in silico* tools be helpful in deciphering the coding sequences of regulatory enzymes involved in the key metabolic pathways (b) how the expression profiles of important genes of key pathways or whole genome network of a crop get changed, when it is grown in stressed soils/environments, or when crops come in contact with pathogens/beneficial microorganisms (c) what changes in the amino acids during divergent evolution led to the increased pathogenicity of microorganisms against agricultural crops? (d) how the carbon fluxes or biomass production get changed with the inoculation of stress mitigating plant growth promoting rhizobacteria (PGPRs) (e) based on *in silico* results, could it be possible to *in vitro* optimize the sequence/codon of a stress mitigating enzyme through modern approaches of synthetic biology and genetic engineering to get the improved crop production/varieties?