JFL

Journal of Food Legumes 35(3): 161-162, 2022

Current Affairs

Molecular insights into the genome of an arid legume: Clusterbean (*Cyamopsis tetragonoloba*)

Dr Kishore Gaikwad

Principal Scientist
National Institute of plant Biotechnology,
LBS Building, New Delhi

Email: *kish2012@gmail.com*

Dr. Kishor Gaikwad, Ph.D. (Agricultural Biotechnology), is a molecular biologist and biotechnologist whose major areas of research revolve around the utilization of advanced molecular biology and breeding tools in the elucidation of genes and pathways regulating vital traits, with prime emphasis on *Cajanus cajan, Vigna mungo* and *Cyamopsis tetragonoloba*. With over two decades of teaching and research experience in genomics and over 100 research papers under his name, Dr. Gaikwad has been a vital group member of several genome sequencing projects like the Rice Genome Sequencing Consortium, Tomato Genome Sequencing Consortium, and Pigeonpea genome sequencing team.

Legumes are a vital component of agriculture as well as global economy. Most of the major grain species of the subcontinent have limited genomic resources *vis a vis* a rice or wheat genomics data. This has to some extent limited the genetic improvement in the new genomics era. Among these major grain legumes, soybean was the first leguminous crop to be sequenced (Schmutz *et al.,* 2010). This was followed by pigeonpea (Varshney *et al.,* 2012 & Singh *et al.,* 2012), chickpea (Varshney *et al.,* 2013), common bean (Schmutz *et al.,* 2014), mung bean (Kang *et al.,* 2014), adzuki bean (Kang *et al.,* 2015), peanut (Bertioli *et al.,* 2015) and cowpea (Lonardi *et al.,* 2019) which were sequenced and assembled to draft levels and improved versions came thereafter. Once the entire genome sequence at chromosomal scale is available, it aids to generate a wealth of useful information including number of genes, repeat sequences and their positions, copy number variation of a gene, and so on.

The current understanding of the cluster bean genome is the first among galactomannan (gum) generating plants. Clusterbean (2n=14), also known as guar, is one of the underutilized legume crops having immense economic importance. It is a climate-resilient annual legume bearing considerable potential as an alternative crop in arid and semi-arid environments. Only one of the four species in the *Cyamopsis* genus, *C. tetragonoloba* (L.) Taub, is cultivated. The other two species, *C. serrata* Schinz and *C. senegalensis* Guill & Perr, are its wild relatives, whereas *C. dentate* Tarre is cosidered to be evolved as a result of an inter-species hybridization between *C. serrata* and *C. senegalensis*. About 80-90% of the endosperm is made up of guar gum (or) galactomannan, a very viscous water-soluble hetero-polysaccharide with a galactose-to-mannose ratio of 1:2. Besides being a rich source of commercial products such as gum, clusterbean is also a highly nutritious legume crop, with a high protein content (18%) and dietary fiber (32%).

A significant portion of genome-level studies have been conducted, mostly by our group, including studies on genetic diversity, genes and pathways governing gum biosynthesis, organelle genomics etc. Simultaneously, we have been working towards developing a high-quality reference genome of the well-known RGC-936 (known for its high yield and high galactomannan content) clusterbean cultivar. Using a hybrid assembly approach to assemble the sequencing data, we could assemble the entire 550Mb of the guar genome which were further anchored to the genetic map with 7 linkage groups.

A total of >34000 protein coding genes were predicted, of which 28000 (78.93%) were successfully annotated in the genome. The genome assembly has high BUSCO scores and LAI index. This high-quality reference genome of clusterbean enabled us to perform comparative analysis among some of the representative plant genomes. Through comparative genomics, we conducted a gene family clustering analysis using cluster bean and 11 other

representative angiosperm species of crop and model plants. The results indicated that clusterbean and other legumes may have shared a common ancestor ~80 mya.

The "-omics" era has given rise to a new set of tools and approaches that are having a substantial influence on the advancement of genetic engineering, metabolic engineering, and synthetic biology. This has sped up the search for the genes and enzymes of specific pathways that are used in metabolic engineering applications and has helped to build the metabolic pathways, which in turn has assisted in the creation of the desired molecules. The current understanding of the cluster bean genome is the first among galactomannan (gum) generating plants. As a result, our research gives valuable information that might be used in future for clusterbean enhancement.