

Decoding the genomes of plants uniquely important for India

Nagendra K. Singh

National Research Centre on Plant Biotechnology, Pusa Campus, New Delhi-110012

Email: nksingh4@gmail.com

After missing the boat on the pathbreaking international initiatives for decoding the genomes of human and *Arabidopsis*, India started its genomics journey by joining the international rice genome sequencing project and sequenced long arm of chromosome 11 (IRGSP, *Nature* 11 Aug 2005). With the experience gained through rice genome we became a leading partner in the international tomato genome sequencing consortium (ITGSC, *Nature* 31 May 2012) and are presently engaged in the decoding of wheat genome also through international partnership (IWGSC, *Science* 18 Jul 2014). While rice, tomato and wheat are crops of global importance which made it feasible to develop international consortia with funding supports from the respective governments, there are many tropical crops, horticultural and medicinal plants uniquely important to India with large production volume where no other nation has high stake, and therefore development of genomic resources in such crops is a priority for India that will pay rich dividend by application in genetic improvement and productivity enhancement. Keeping this in mind we started work on development of genomic resources in pigeonpea, jute and mango by forging national partnerships with funding support from ICAR. Pigeonpea is the second most important pulse legume for India after chickpea and is grown in all parts of the country. India is by far the largest producer, consumer and even importer of pigeonpea. It has a genome size of 858 Mbp packed in its 11 chromosomes. After developing huge amount of RNA sequence data and transcriptome sequence assemblies we have created high density genetic maps of pigeonpea and the first draft of the pigeonpea genome in 2011. This was the first plant genome decoded entirely by indigenous efforts and first genome of a pulse crop internationally. The pigeonpea genome was predicted to have 47,004 protein coding genes out of which 1213 were disease resistance and defence response like genes. A large number of SSR and SNP markers have been identified two major pleiotropic QTLs for semi-dwarf plant height, early maturity and determinate growth habit were mapped and are now being used in the pigeonpea breeding programme. A 62K SNP genotyping chip has also been developed on Affymetrix platform for high density genotyping. Mango is one of the largest tropical fruit crops of the world in terms of production volume and India is the largest producer with growing international demand for high quality fruits, but no genomic resources were available till recently. A major challenge for mango genome assembly is its high level of heterozygosity and lack of open source software for diploid genome assembly. To solve this problem we used long sequence reads PacBio single molecular real time (SMRT) sequencing and Falcon diploid genome assembler to generate a 403 Mb of high quality reference genome of the hybrid mango variety Amrapali and predicted 46,283 protein coding genes. It is expected that development of genomic resources and NGS genotyping tools will speed up genetic improvement of varieties in rice, pigeonpea, jute and mango for productivity enhancement, quality and stress tolerance. India has a large number horticultural, medicinal and aromatic plants with high economic values for which genomic resources are meager. Hence, there is need to take of genome sequencing and develop high density fingerprint of these for effective crop improvement and monitoring of planting materials.