Agri Articles
(e-Magazine for Agricultural Articles)
Volume: 02, Issue: 04 (JULY-AUGUST, 2022)
Available online at http://www.agriarticles.com
©Agri Articles, ISSN: 2582-9882

Application of Next Generation Sequencing Technology in Agriculture

(*R.R.Vaid)
C. P. College of Agriculture, Sardarkrushinagar Dantiwada Agricultural University, Sardarkrushinagar - 385 506, Gujarat, India
*Corresponding Author's email: rahulpatel5767@gmail.com

The development of DNA sequencing strategies has been a high priority in genetics research since the discovery of the structure of DNA and the basic molecular mechanisms of heredity. However, it was not until the works by Maxam and Gilbert (1977), and Sanger (Sanger *et al*, 1977), that the first practical sequencing methods were developed and implemented on a large scale. The first isolation and sequencing of a plant cDNA by Bedbrook *et al.* a few years later initiated the field of Plant Molecular Genetics. Plant biotechnology started shortly thereafter with the successful integration of recombinant DNA and sequencing techniques to generate the first transgenic plants using *Agrobacterium*.

The determination of the reference genomes in *Arabidopsis thaliana*, rice and maize using Sanger sequencing strategies constituted major milestones that enabled the analysis of genome architecture and gene characterization in plants by The Arabidopsis Genome Initiative in 2001 and International Rice Genome Project in 2005(Schnable *et al*, 2009). More recently, the development and increasing availability of multiple Next-Generation sequencing (NGS) technologies minimized research limitations and bottlenecks based on sequence information (Metzker, 2010; Glenn, 2011). It is difficult to overstate the influence that these massively parallel systems have had in our understanding of plant genomes and in the expansion, acceleration and diversification of breeding and biotechnology projects. At the same time, this influence tends to understate the importance that capillary Sanger sequencing still has in day-by-day research and development work.

DNA sequencing

The determination of base sequence of a DNA fragment is called **DNA sequencing**. Initially two methods of DNA sequencing were developed, chemical method and enzymatic method by Allen Maxam Walter Gilbert and Fred Sanger in 1977 respectively.

Chemical method uses specific chemical modifications of DNA bases, ultimately, leading to breaks in DNA strands at the sites occupied by the modified bases. Four separate reactions are set up for the modification of different bases, and gel electrophoresis, followed by autoradiography.

Enzymatic method is based on the selective incorporation of chain-terminating dideoxy-nucleotides by DNA polymerase during *in vitro* DNA replication in a four separate reactions.

Next generation sequencing method

The next-generation DNA sequencing methods, also called massively parallel sequencing (MPS) technologies are faster, cheaper and require much less template preparation than the Sanger–Coulson method. The NGS methods use PCR amplification for template preparation. They allow simultaneous sequencing of hundreds of thousands to hundreds of millions of different DNA fragments.

NGS Platforms

Different NGS platforms use different technologies for template preparation for example **The 454 platform/ Roche 454, Solexa platform/ Illumina** and **Ion torrent** use sequencing by synthesis method while **SOLiD** of life technologies and **polonator** uses sequencing by ligation method for DNA sequencing.

Application of Next Generation Sequencing Technology

(i) Full-genome resequencing or more targeted discovery of mutations or polymorphisms.

(ii) Mapping of structural rearrangements, which may include copy number variation, balanced translocation breakpoints and chromosomal inversions.

(iii) 'RNA-Seq', analogous to expressed sequence tags (EST) or serial analysis of gene expression (SAGE), where shotgun libraries derived from mRNA or small RNAs are deeply sequenced.

(iv) Large-scale analysis of DNA methylation, by deep sequencing of bisulfite-treated DNA.

(v) 'ChIP-Seq', or genome-wide mapping of DNA-protein interactions, by deep sequencing of DNA fragments pulled down by chromatin immuno precipitation.

Review of literature

Causse *et al.* (2013) sequenced eight genome using the GAII Illumina platform. Comparison of the sequences with the reference genome yielded more than 4 million single nucleotide polymorphisms (SNPs). Almost 128,000 InDels were detected. The distribution of SNPs and InDels across and within chromosomes was highly heterogeneous revealing introgressions from wild species. In-depth annotation of the polymorphisms identified more than 16,000 unique non-synonymous SNPs. In addition 1,686 putative copy-number variations (CNVs) were identified.

Zamora *et al.* (2013) studied 5528 SNPs of which 1980 originated from 454-sequencing, 3495 from Illumina Solexa sequencing and 53 were additional known markers. Genotyping different tomato samples allowed the evaluation of the level of heterozygosity and introgressions among commercial varieties and concluded that Cherry tomatoes were especially different from round/beefs in chromosomes 4, 5 and 12. They were able to identify a set of 750 unique markers distinguishing *S. lycopersicum* 'Moneymaker' from all its distantly related wild relatives. Clustering and neighbor joining analysis among varieties and species showed expected grouping patterns, with *S. pimpinellifolium* as the most closely related to commercial tomatoes earlier results.

Ashelford *et al*. (2011) Re-sequenced the 120-Mb genome of a novel *Arabidopsis* clock mutant early bird (ebi-1) in Wassilewskija (Ws-2). Demonstrated the utility of sequencing a backcrossed line in limiting the number of SNPs considered and identified a SNP in the gene *AtNFXL-2* as the likely cause of the ebi-1 phenotype.

Bhattacharyya *et al.* (2013) detected 961 transcripts containing transcription factor domains. High performance liquid chromatography analysis showed the peak accumulation of podophyllotoxin in 12-day cell suspension cultures. A comparative qRT-PCR analysis of phenylpropanoid pathway genes identified in the present data as performed to analyze their expression patterns in 12-day cell culture, callus and rhizome.

Prajapat *et al.* (2017) identified the regulated transcription factor, genes and biochemical pathways that impart tolerance against draught stress condition by using illumina Hiseq 2500 sequencer. High quality reads of all the cotton leaf sample were mapped on the reference genome *Gossypium hirsutum L.* and estimated gene expression using different bioinformatics tools.

Conclusion

Advancement in sequencing technologies has had a great impact on crop genetics, enabling the sequencing of genomes and transcriptomes of several crops. Although, reference genomes have been obtained for many important crops, massive re-sequencing and gene expression studies are essential to identify the key genes responsible for a desired trait and to find its allele variability. Utilization of this knowledge in crop breeding would empower the development of better crop varieties and may lead to a second green revolution. This would reduce the hunger of billions and revolutionise the economies of developing countries.

References

1. Ashelford, K., Eriksson, M., Allen, C., D'Amore, R., Johansson, M., Gould, P., Kay, S., Millar, A., Hall, N. and Hall, A. (2011). *Genome Biology*, **12**:R28.
2. Bhattacharyya, D., Sinha, R., Hazra, S., Datta, R. and Chattopadhyay, S. (2013) *BMC Genomics*, **14**:748.
3. Causse, M., Desplat, N., Pascual, L., Paslier, M., Sauvage, C. and Bounon, R. (2013) *BMC Genomics*, **14**:791
4. Glenn, T.C. (2011). *Molecular Ecology Resources* **11**: 759-769.
5. Maxam, A.M. & Gilbert, W. (1977). *Proceedings of the national academy of Sciences*, USA **74**: 560–564.
6. Metzker, M.L. (2010). *Nature Reviews* **11**:31-46.
7. Prajapat *et al.* (2017) In*: National conference on breaking barriers through bioinformatics and computational biology*. PP 109.
8. Sanger, F., Nicklen, S. and Coulson, R. (1977). *Proceedings of the national academy of Sciences*, USA **74**:5463-5467.
9. Schnable, P.S., Radovic, S., Rafalski, A. (2009).*Science* **326**: 1112-1115.
10. Zamora, M., Vosman, B., Geest, H., Bovy, A., Visser, R., Finkers, R. and Heusden, A. (2013). *BMC Genomics*, **14**:354.