Future Prospects for Peanut Improvement

11

Rajeev K. Varshney, Manish K. Pandey and Naveen Puppala

Abstract

Evolution in sequencing technologies led to reduction in costs and increase in speed for generating sequence data. The affordability of low-cost sequencing is expected to make other genotyping platforms obsolete in next couple of years. The concept of "single genome sequence" in a crop has evolved to sequencing of multiple genomes to assemble pangenomes. Sequencing combined with precise phenotyping of segregating populations and germplasm collections is expected to measure the accurate genetic diversity present in the germplasm as well as to identify the gene/nucleotide associated with the trait(s). It is time now to move toward using multi-parents populations from bi-parental populations for trait discovery and identify superior haplotypes. Availability of information on functional variation for genes controlling traits of interest will eventually help in manipulating genes more routinely using appropriate technologies such as marker-assisted selection/backcrossing, genomic selection, and genome editing. This chapter provides expected use of genome sequence and allied information on peanut for accelerating biology research as well as peanut improvement.

11.1 Introduction

The genus *Arachis,* containing 81 species arranged in nine taxonomic sections and variable genomes, face a huge genetic barrier for bringing exotic alleles to cultivated species from wild relatives (Pandey et al. 2012). Such genetic barrier not only limits the enhancing of genetic variation between different genepools, but also presents the challenge to researchers working on wild relatives for trait discovery and their

R.K. Varshney (✉) · M.K. Pandey
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India
e-mail: r.k.varshney@cgiar.org

R.K. Varshney
University of Western Australia (UWA), Crawley, Australia

N. Puppala
New Mexico State University, Agricultural Science Center, Clovis, USA

R.K. Varshney et al. (eds.), *The Peanut Genome*, Compendium of Plant Genomes, DOI 10.1007/978-3-319-63935-2_11

deployment in breeding improved varieties. The existing trait dissection and breeding methodologies offer limited scope for large-scale modifications of genetic composition of lines. Therefore, advanced technologies must be developed and deployed to understand the complex and multiple genomes exist in genus *Arachis*. This will help in devising better and improved technologies and strategies for discovery of functional sequence variations in the genome associated with desirable traits in peanut. Availability of such information on desirable traits of peanut will facilitate faster alterations of multiple genes in the genome using modern molecular breeding technologies including genome editing.

The last decade has witnessed extraordinary progress in genome sequencing technologies leading to faster speed and huge reduction in cost per megabase. The high-throughput DNA sequencing technologies have achieved unprecedented scale of efficiency in sequencing followed by improved analytical tools for analyzing large-scale sequencing data to understand the complex biological problems. As a result, it has been possible to sequence several crop genomes including complex genomes (Goodwin et al. 2016). Genome sequence has become possible for AA and BB genome progenitor species of peanut (Bertioli et al. 2016; Chen et al. 2016). The onus, therefore, lies now on the peanut researchers for enhancing understanding of different *Arachis* genomes, identifying functional sequence variation followed by their deployment in peanut breeding using appropriate technologies and methods. Editors highlight some areas that can be addressed in coming years for accelerating peanut genetics and crop improvement.

11.2 Breaking Species/Section Barriers for Enhancing Genetic Base

Arachis genus is arranged in nine taxonomic sections carrying different genomes namely *Arachis* (A, B, D, F, G, and K), Trierectoides (E), Erectoides (E), Extranervosae (Ex), Triseminatae (T), Heteranthae (Am), Caulorrhizae (C), Procumbentes (E), and Rhizomatosae (R) (Krapovickas and Gregory 1994; Valls and Simpson 2005). Further, the cultivated peanut (*Arachis hypogaea* L.) can be divided into two subspecies (*hypogaea* and *fastigata*) based on the morpholocal differences in branching pattern and vegetative and reproductive axes. Based on inflorescence, pod, and seed characters, these subspecies can be further subclassified into six botanical varieties, i.e., *fastigata*, *vulgaris*, *peruviana*, and *aequatoriana* from subspecies *fastigata,* while *hypogaea* and *hirsuta* from subspecies *hypogaea.* In addition to above classifications, the peanut is also divided into different kinds of market types based on the seed size, plant growth type, and its applications in peanut industry, i.e., Spanish, Runner, Virginia, and Valencia. Currently, most of the breeding programs across the globe are engaged in developing improved varieties for one or two botanical and market types based on the demand in local and international markets. Genomics can play a major role in developing better understanding on different genomes of wild relatives, botanical, and market types so that improved varieties with specific features suitable to specific climatic conditions can be developed using genome-based breeding approaches.

11.3 Sequencing Reference Genome and Germplasm Collection for Developing Pangenomes and Hapmap

The cultivated peanut (*A. hypogaea*) is an allotetraploid ($2n = 4x = 40$) crop with two subgenomes. The Peanut Genome Consortium (PGC) with the collaboration of international partners developed first draft sequences of two progenitors of tetraploid cultivated peanut, representing A-genome (*Arachis duranensis*, accession V14167) and B-genome (*A. ipaensis,* accession K30076) (Bertioli et al. 2016). Another consortium namely Diploid Progenitor A-genome Sequencing Consortium (DPPAGSC) developed another draft sequence of A-genome

progenitor (*A. duranensis*, accession PI475845) (Chen et al. 2016). The effort by PGC is continued to develop a very high-quality tetraploid genome sequence for cultivated peanut by the end of this year.

Despite several efforts using cytogenetic and genetic studies, the level of genome diversity, genome evolution, and accurate phylogenetic relationship could not be established with a high level of precision and confidence in *Arachis* genus. Sequencing of mere two diploid progenitors and cultivated tetraploid genotypes do not represent the sequence variation present in the entire germplasm. It will be good to sequence all the genomes available in *Arachis* species including representative genotypes from each species from different sections. Although >15,000 accessions available in different genebanks at ICRISAT, USDA/ARS-Griffin, Georgia, USA, three important minicore collections namely ICRISAT MiniCore Collection (184 accessions), Chinese MiniCore Collection (298 accessions), and US MiniCore Collection (112 accessions) as well as three core collections namely ICRISAT Core Collection (1704 accessions), US Core Collection (831 accessions), and Chinese Core Collection (576 accessions) in addition to global Composite Collection (1000 accessions) can be started for resequencing in systematic manner. These efforts will provide core genome and pangenome helping in understanding genome evolution in a better way leading to answer key questions related to genome variations, evolution, phylogenetic relationship, and potential method of exchanging genome variations across *Arachis* genus. It will also be desirable to undertake phenotyping of those lines so that genome-wide association study (GWAS) at high-resolution level can be undertaken. Such an analysis will help to identify genes for traits of interest, superior haplotypes for a given gene (associated with traits of interest) and lines comprising suitable haplotypes for different genes. As the sequencing cost is reducing day by day, we are hopeful that global germplasm collections including in peanut will be sequenced in the long run.

11.4 Sequencing-Based Trait Dissection and Gene Discovery

In most of the current trait mapping studies, high-throughput genotyping using different kinds of SNP genotyping platforms is used for conducting linkage or GWAS studies. It has become possible now to use sequencing-based genotyping of the segregating populations. We anticipate that in the coming years, sequencing will be the approach of genotyping of mapping populations/germplasm sets for conducting high-resolution mapping and faster discovery of candidate genes for developing diagnostic markers for traits of interest. Most importantly, several analytical softwares have also become available for analyzing large datasets (Varshney et al. 2015). At present also, genotyping-by-sequencing (GBS), skim sequencing, and BSA-Seq/QTL-Seq approaches have been used for trait mapping (Pandey et al. 2016).

For high-resolution mapping, it is also essential to start using multi-parent genetic populations such as multi-parent advanced generation intercross (MAGIC), nested association mapping (NAM), and recombinant inbred advanced intercross line (RIAIL) populations (Morrell et al. 2012; Pandey et al. 2016). ICRISAT has developed three specialized MAGIC populations for aflatoxin resistance, drought tolerance, and nutritional and quality traits. Similarly, one NAM population each has been developed for Spanish and Virginia types by ICRISAT. These two types of multi-parent populations also provide opportunity to conduct joint linkage-association mapping (JLAM) in addition to linkage mapping. Such complex genetic populations often have several hundreds of individuals for genotyping and phenotyping similar to the majority of the association mapping panels. The low-cost sequencing will allow researchers in coming years to perform sequencing of complete mapping population/panel for conducting high-resolution trait mapping and candidate gene discovery. These developments will help in dissecting even the most complex traits such as drought tolerance, aflatoxin contamination, and allergens.

11.5 Next-Generation Breeding

The current molecular breeding approaches deploy genetic marker information using three major approaches namely marker-assisted backcrossing (MABC), marker-assistant recurrent selection (MARS), or genomic selection (GS) in crops including peanut (Varshney et al. 2013; Pandey et al. 2016). The successful example can be cited for improving selected traits such root-knot nematode (Chu et al. 2011), high oleic acid (Chu et al. 2011; Janila et al. 2016), and foliar fungal diseases (Varshney et al. 2014). However, there is no report available on deployment of other two approaches (MARS and GS) in peanut. Nevertheless, ICRISAT has taken some initiatives toward developing training population and completed its genotyping using 58 K SNP array developed recently (Pandey et al. 2017). Phenotyping of this training population is underway for different key agronomic traits to develop genomic selection model and initiation of GS breeding for some selected traits.

The other breeding approach is early generation screening (EGS) of large populations with at least markers for must-have traits. This will help enhancing selection intensity and in turn accelerate genetic gains in the breeding program. However, to deploy the EGS approach, it is essential to have diagnostic markers for majority/must-have traits.

As a result of large-scale genome resequencing projects, it will be possible to identify not just causal gene but also causal nucleotide for a given trait. In that scenario, it will be possible to undertake genome editing approach (Wood et al. 2011). In addition to adding the favorable alleles, genome editing also offers removal of deleterious alleles that have become available as a result of accelerated domestication of wild relatives. We envisage use of combination of EGS, GS, and genome editing in peanut in coming years. Of course, MABC should be continued to improve the elite/mega varieties for 1–2 traits for which varieties are deficient.

11.6 Conclusion

NGS technologies have undoubtedly accelerated the genomics research drastically leading to the generation of large data at reduced costs and less time. Sequencing of entire set of genomes, species, botanical varieties, and genbank germplasm will improve current understanding to devise novel strategies for harnessing the sequence diversity present across genepools. The low-cost sequencing will allow sequencing-based genotyping of large-scale populations/panels containing thousands of individuals for conducting high-resolution trait dissection and gene discovery, thereby making trait mapping more reliable and less time consuming. Given the speed of evolution in NGS technologies and advances in developing decision support tool, the next-generation breeding approaches will be used for peanut improvement.

References

Bertioli DJ, Cannon SB, Froenicke L et al (2016) The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nat Genet 48:438–446

Chen X, Li H, Pandey MK et al (2016) Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens. PNAS 113:6785–6790

Chu Y, Wu CL, Holbrook CC, Tillman BL, Person G, Ozias-Akins P (2011) Marker assisted selection to pyramid nematode resistance and the high oleic trait in peanut. Plant Genome 4:110–117

Goodwin S, PcPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17:333–351

Janila P, Pandey MK, Shasidhar Y et al (2016) Molecular breeding for introgression of fatty acid desaturase mutant alleles (ahFAD2A and ahFAD2B) enhances oil quality in high and low oil containing peanut genotypes. Plant Sci 242:203–213

Krapovickas A, Gregory WC (1994) Taxonomía del género *Arachis* (Leguminosae). Bonplandia 8:1–186

Morrell PL, Buckler ES, Ross-Ibarra J (2012) Crop genomics: advances and applications. Nat Rev Genet 13:85–96

Pandey MK, Monyo E, Ozias-Akins P, et al (2012) Advances in Arachis genomics for peanut improvement. Biotechnol Adv 30(3):639–651

Pandey MK, Agarwal G, Kale SM, et al (2017) Development and evaluation of a high density genotyping 'Axiom_*Arachis*' array with 58 K SNPs for accelerating genetics and breeding in groundnut. Sci Rep 7:40577

Pandey MK, Roorkiwal M, Singh VK et al (2016) Emerging genomic tools for legume breeding: current status and future prospects. Front Plant Sci 7:455

Valls JFM, Simpson CE (2005) New species of *Arachis* (Leguminosae) from Brazil, Paraguay and Bolivia. Bonplandia 14:35–63

Varshney RK, Mohan SM, Gaur PM et al (2013) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. Biotechnol Adv 31:1120–1134

Varshney RK, Pandey MK, Janila P et al (2014) Marker assisted introgression of a QTL region to improve rust resistance in three elite and popular varieties of peanut (*Arachis hypogaea* L.). Theor Appl Genet 127:1771–1781

Varshney RK, Singh VK, Hickey J et al (2015) Analytical and decision support tools for genomics-assisted breeding. Trends Plant Sci. doi:10.1016/j.tplants.2015.10.018

Wood AJ, Lo T-W, Zeitler B, et al (2011) Targeted genome editing across species using ZFNs and TALENs. Science (New York, NY) 333(6040):307