Himachal Journal of Agricultural Research 44(1&2) : 125-127 (2018)

Short Communication

Emergence of genome wide association studies as a major plant breeding tool

Neha Sharma*, R.K. Mittal, V.K. Sood and G. Katna
Department of Crop Improvement
CSK Himachal Pradesh Krishi Vishvavidyalaya, Palampur-176 062, India.
*Corresponding author: neha.sh914@gmail.com

Manuscript Received: 05.01.2018; Accepted: 17.02.2018

Abstract

Identification of genes responsible for complex traits has been revolutionized with the advent and association of high throughput sequencing with association studies. GWAS takes the full advantage of ancient recombination events occurred in the natural population and exploits simple association of polymorphism with the trait of interest. Further, the genetic basis of phenotypic variations is detected by rapid high-throughput sequencing, accelerating the genome assisted breeding. The high density SNPs generated through GWAS allow whole genome scanning to identify haplotype blocks correlated with desired QTLs, thereby establishing itself as a major plant breeding tool.

Key words: GWAS, Linkage disequilibrium, SNPs.

Genome wide association studies (GWAS) has vindicated itself not only in the identification of variation associated with human genetic disorders but also in plant science for various agronomic traits, abiotic stress resistance and biotic disease resistance. This is a "phenotype first" approach in which the unrelated population is first screened for the presence or absence of the phenotype. It further involves rapid screening of complete genome of diverse population to find maximum number of genetic variations associated with a particular trait. While a linkage study relies on co-segregation of gene and the marker, GWAS exploits simple association of polymorphism (SNP) with the trait of interest. The size of the population is advised to be fairly large so that association between SNP and the trait of interest is validated. If certain variation is significantly found in population with the concerned trait, the variation is said to be associated with the trait. Entire genome is screened for polymorphism rather than the small number of pre-specified genetic regions. GWAS is also known as linkage disequilibrium (LD) analysis that represents the lack of recombination between the neighbouring SNPs. The set of such SNPs associated, are called haplotypes. The haplotype maps of large population are useful to identify rare alleles. Rice diversity project aimed to develop 10,000 SNP chips for *O. sativa* and *O. rufipogon* populations (Anonymous, 2018), created a haplotype document for species. In legumes, 30 inbred lines of *Medicago* were sequenced to explore the genetic basis of symbiosis (Branca, 2010). The major problem with the association studies is the LD decay over time as recombination breaks haplotypes. To prevent this, the genetic diversity of the population must first be analysed to assess the fitness of population for association studies. Being a non candidate approach, it only specifies the variation associated with the trait and not the candidate gene behind. GWAS is mostly performed on unrelated diverse germplasm to maximize the diversity captured but the QTLs identified must be validated before they can be used in genome assisted selection. For such validation, it is important to perform GWAS on adapted lines so that the QTLs can be directly used in marker assisted selection in breeding programmes. The alliance of GWAS with the modern day gene mapping technologies has revolutionised the candidate gene identification. Han *et al.* (2018) identified five candidate genes for the capsaicinoid content in pepper using two RIL populations and one GWAS population. Mapping of HFT1; 5 gene in barley by Hazzouri *et al.*

(2018) and identification of the underlined molecular mechanism of the gene in salt tolerance in barley clearly demonstrates the power of the GWAS technology when merged with other high throughput cloning and sequencing technologies.

GWAS has emerged as an efficacious tool for the high throughput genotyping. Sanchez *et al.* (2018) found that SNP S5_152926936 in chromosome 5 of maize is associated with seedling root development. The study involved the GWAS of 300 doubled haploid lines derived from cross between GEM (Genetic enhancement of maize) accessions and inbred lines PHB47 and PHZ51. Avirulence factor AvrStb 6 of *Zymoseptoria tritici*, recognized by Stb 6 resistance gene in wheat was identified through GWAS and classic linkage mapping by Zhong *et al.* (2017) for effector discovery in the pathosystems of plants which had been a challenging task. In the diverse collection of 421 Chinese soybean, using 1536 SNPs, Li *et al.* (2018) identified six SNPs associated with seed oil and protein content each and two controlling the oil content, by integrating the results of association mapping with known QTLs and tissue specific data. Li *et al.* (2018) revealed the genetic basis of correlation among growth duration and yield traits in rice through GWAS. The 47 pleiotropic genes in 15 pleiotropic QTLs were detected for the studied traits and the reason of correlation was revealed to be pGenes and linkage between two genes controlling different traits along with the interaction between genes within the same biological pathway. Wilson *et al.* (2018) studied the population structure of *Brachypodium* species, a model plant, in response to climate through GWAS across simulated seasonal environments and found a total of 8, 22 and 47 QTLs for flowering time, early vigour and energy traits, respectively. Batayeva *et al.* (2018) found 26 QTLs associated with salinity stress tolerance, 11 of which were co-located with salinity tolerance gene. These QTLs and gene clusters were found on chromosome 1, 2, 3, 6, 8 and 9. Han *et al.* (2018) investigated the genetic architecture of Gibberella ear rot in maize with GWAS within two elite panels of 130 dent and 114 flint lines and reported low genomic prediction among the pools.

Main challenge in agriculture is to use tremendous genetic variation present in available germplasm which remains untapped due to difficulty in effectively identifying genetic differences. Traits having high heritability and easy characterization can be easily selected while the ones with complex inheritance are tedious to select. For such complex traits, GWAS is capable of capturing the maximum variability. GWAS, a well established methodology in human genetics, has successfully established itself as a major plant breeding tool. The DNA markers associated with the trait identified can be used in MAS (marker assisted selection) for the incorporation of the desirable alleles. With the emergence of low cost high throughput sequencing technologies, GWAS has come up as a cost effective and emerging tool for genotyping and assisting fine mapping, thereby increasing the scope of this methodology in crop improvement.

References

Anonymous. 2016. http://www.ricehapmap.org/index.aspx.

Branca A, Paape T, Briskine R, Zhou P, Wang S, Denny R, Mudge J, Bharti AK, Farmer A, May GD, Tiffin PL and Young ND. 2010. The *Medicago truncatula* HapMap project: deep coverage sequencing of 30 inbred lines using Illumina's Solexa technology. Plant & Animal Genomes XVIII Conference, San Diego, CA pp: 417.

Hazzouri KM, Khraiwesh B, Amiri KMA, Pauli D, Blake T, Shahid M, Mullath SK, Nelson D, Mansour A, Ashtiani KS, Purugganan M and Masmoudi K. 2018 . Mapping of HKT1;5 gene in barley using GWAS approach and its implications in salt tolerance mechanism. *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2018.00156.

Han K, Lee, HY, Ro NY, Hur OS, Lee JH, Kwon JK and Kang BC. 2018. QTL mapping and GWAS reveal candidate gene controlling capsaicinoid content in Capsicum. Plant Biotechnology. https:// doi.org/ 10.1111/pbi.12894.

Batayeva D, Labaco B, Ye C, Li X, Usenbekov B, Rysbekova A, Dyuskalieva G, Vergara G, Reinke R and Leung H. 2018. Genome wide association study of seedling stage salanity tolerance in temperate japonica rice germplasm. BMC Genetics **19**:2

Wilson PB, Streich JC, Murray KD, Eichten SR, Cheng R, Aitken NC, Spokas K, Warthmann N and Borevitz JO. 2018. Population structure of the *Brachypodium* species complex and genome wide association of agronomic traits in response to climate. *bioRxiv*. https:// doi.org/ 10.1101/246074.

Sanchez DL, Liu Sisi, Ibrahim R, Blanco M and Lubberstedt T. 2018. Genome-wide association studies of doubled haploid exotic introgression lines for root system architecture traits in maize (*Zea mays* L.). Plant Science **268**: 30-38.

Zhong Z, Marcel TC, Hartmann FE, Ma X, Plissonneau C, Zala M, Ducasse A, Confais J, Compain J, Lapalu N and Amselem J. 2017. A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the Stb6 resistance gene. New Phytologist **214**: 619-631.

Li YH, Reif JC, Hong HL, Li HH, Liu ZX, Ma YS, Li J, Tian Y, Li YF, Li WB and Qiu LJ. 2018. Genome wide association mapping of QTS underlying seed oil and protein contents of a diverse panel of soybean accessions. Plant Science **266**: 95-101.

Li F, Xie J, Zhu X, Wang X, Zhao Y, Ma X, Zhang Z, Rashid MAR, Zhang Z, Zhi L, Zhi L, Zhang S, Li J, Li Z and Znag H. 2018. Genetic basis underlying correlations among growth duration and yield traits revealed by GWAS in rice (*Oryza sativa* L.). Frontiers in Plant Science. https://doi.org/10.3389/fpls.2018.00650.