

Genome-Wide Association Studies (GWAS): Fast Forward Approach for Gene Mapping

Lavale S. A., Kelkar V.G., and Feba Jacob

Ph.D., Research Scholar, Centre for Plant Biotechnology and Molecular Biology, Kerala Agricultural University, Thrissur, Kerala

SUMMARY

Identification of genes/QTLs responsible for a complex trait is a crucial step in plant breeding. With the advent in molecular marker technology QTL mapping has become an easy affair. However, the time taken for identification of these QTLs has direct impact on flow of plant breeding. Conventional biparental mapping technology, though reliable, takes very long time to generate mapping population and mapping the trait. It also finds the QTLs with lower resolution. Association mapping technology, on the other hand, is getting proved as fastest way to map the complex traits. It can identify QTLs with high resolution, which is attributed to use of diverse panel of germplasm for mapping. Most importantly it can be utilized in perennial plant species.

INTRODUCTION

In plants, most of the QTL analysis have been conducted using highly structured populations with known pedigrees. However, such structured populations have limited number of recombination events resulting poor resolution for quantitative traits, and are capable of studying only two alleles at any given locus. Hence, large populations which have undergone several rounds of random mating should be created to increase the resolution of mapping (Boopathi, 2020). Association analysis, also known as linkage disequilibrium (LD) mapping or association mapping (AM), is a population-based survey used to identify the trait-marker relationships. Unlike linkage analysis where the familial relationships are used to predict correlations between phenotype and genotype, association analysis rely on previous, unrecorded sources of disequilibrium to create population-wide marker-phenotype associations (Boopathi, 2020). Association mapping is divided into two types: genome-wide association studies (GWAS) and candidate gene (CG) based association. Unlike CG based analysis where preliminary studies are performed and candidate genes are analyzed, GWAS scans the whole genome to determine if any relationship exists between phenotypes and genome-wide markers.

Steps involved in GWAS

The basic steps involved in GWAS are phenotyping, genotyping, population structure (Q) analysis, kinship (K) analysis, and marker trait association analysis (Figure 1).

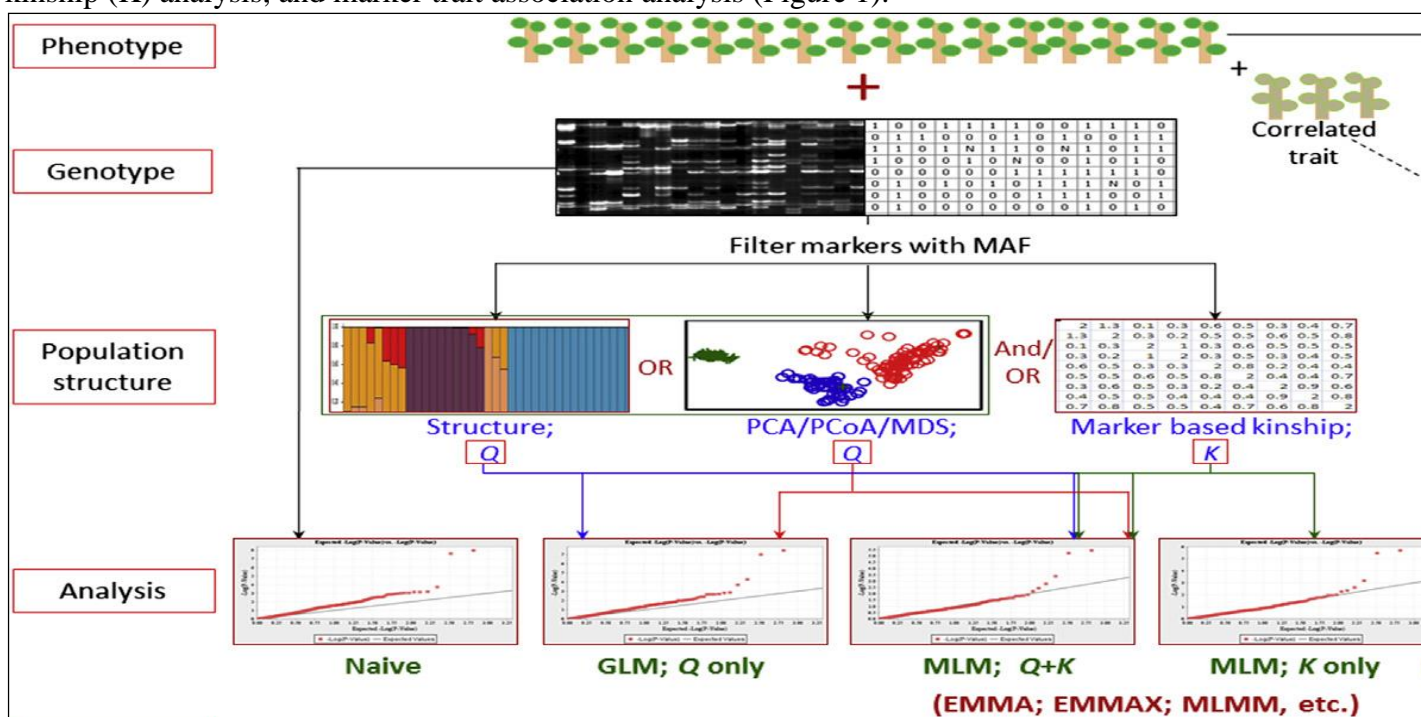


Figure 1. Steps involved in GWAS (Source: Gupta *et al.*, 2014)

Methods used in GWAS

Single locus – Generalized Linear Model (GLM): Uses least squares fixed effects linear model

Single locus – Mixed Linear Model (MLM): Includes both fixed and random effects and corrects for Q (Structure) and/or K (Kinship)

Multi-locus Mixed Linear Model (MMLM): It corrects for Q, K and considers background genotype

Software used in GWAS

TASSEL (Trait Analysis by association, Evolution and Linkage; <http://www.maizegenetics.net/tassel>): Used for marker trait association analysis using SSR/SNP marker genotyping data

STRUCTURE (<http://pritch.bsd.uchicago.edu/structure.html>): Used for analyzing the structure of the population under study SSR/SNP marker genotyping data

GAPIT – R platform: Used for marker trait association analysis

NAM – R platform: Used for marker trait association analysis using genotyping data of NAM population

GWAS is extremely powerful because the individuals tested need not to be closely related, and this enables to utilize all the meiotic and recombination events among the individuals to improve the resolution. Because of these recombination events, only the markers in LD with the trait of interest will get associated with the trait. Nevertheless, more markers are required to cover the genome, considering the anticipated rate of LD decay. With the rapid development of sequencing technologies and computational methods, GWAS is now becoming a powerful tool for detecting natural variation underlying complex traits in crops (Huang and Han, 2014). Most of the GWA studies are done with single nucleotide polymorphisms (SNPs).

CONCLUSION

GWAS helps breeder to efficiently tap the natural variations for unfolding the genomes in very less time. With the advantage of high resolution mapping and inclusion of multiple alleles in GWAS, the limitations of conventional linkage mapping can be effectively resolved.

REFERENCES

- Boopathi, N. M. 2020. Genetic mapping and marker assisted selection: Basics, practice and benefits (2nd Ed.). Springer, Singapore, 504p.
- Gupta, P.K., Kulwal, P.L. and Jaiswal, V., 2014. Association mapping in crop plants: opportunities and challenges. In: Friedmann T., Dunlap J.C. and Goodwin, S.F. (Eds.), *Advances in genetics*, Vol. 85, Academic Press, Waltham, USA, pp. 109-147.
- Huang, X. and Han, B. 2014. Natural variations and genome-wide association studies in crop plants. *Annual Reviews on Plant Biology* 65: 531-551.