Bio. Res. Today 2025, 7(2):41-43

Biotica Research Today

e-ISSN: 2582-6654

February, 2025

Article ID: RT1759

Popular Article

Selective Genotyping: A Rapid and Cost-Effective Approach for QTL Detection

Bidwan Rath[1*], Kavya M.E.[2], Prashantha V.[2], Shirisha K.M.[2] and Kemparaju M.[2]

[1]*Division of Genetics, ICAR-Indian Agricultural Research Institute (IARI), New Delhi (110 012), India*

[2]*Dept. of Genetics and Plant Breeding, University of Agricultural Sciences (UAS), Bangalore, Karnataka (560 065), India*

Open Access

Corresponding Author

Bidwan Rath

✉: bidwan2306@gmail.com

Conflict of interests: The author has declared that no conflict of interest exists.

How to cite this article?

Rath, B., Kavya, M.E., Prashantha, V., *et al*., 2025. Selective Genotyping: A Rapid and Cost-Effective Approach for QTL Detection. *Biotica Research Today* 7(2), 41-43.

Abstract

Detection and mapping of QTLs through the genotyping of an entire mapping population is a time and cost demanding avenue. Genotyping the individuals exhibiting extreme phenotypes in the mapping population for QTL discovery is a cost-effective and rapid alternative. Selective genotyping is one such approach that identifies markers linked to QTLs based on a comparison between marker allele frequencies of individuals showing extreme phenotypes in the mapping population. When the cost of genotyping exceeds that of phenotyping, selective genotyping becomes more cost-effective over conventional QTL mapping. Selective genotyping also offers breeders the opportunity for simultaneous breeding and QTL detection in the segregating generation. Nevertheless, the efficiency of selective genotyping is influenced by various factors that must be optimized prior to its implementation for marker-trait linkage identification.

Keywords: BSA, Genetic hitchhiking, Marker-assisted breeding, QTL

Introduction

Approaches to establish marker-trait associations can be classified into two broad categories. One approach involves genotyping the entire mapping population and classifying the individuals based on marker genotypic classes. A significant difference between marker class genotypes for trait means establishes a marker-trait association. This approach is costly and time-intensive as it involves genotyping the entire mapping population. The other approach targets individuals exhibiting extreme phenotypic values in the mapping population (for example, extremely early- and late-flowering genotypes in a mapping population derived from a cross between early- and late-flowering parents) for genotyping and compares the marker allele frequencies between the two extreme groups. A significant difference in marker allele frequency between the two extreme phenotypic classes establishes marker-trait linkage (Figure 1).

Selective genotyping extends the bulked segregant analysis (BSA) by individually genotyping plants present in the two extreme tails of the phenotypic distribution of a mapping population to identify markers linked to the trait of interest. For large mapping populations, such as those with more than 500 individuals, it facilitates the identification of marker-trait linkage with minimal genotyping effort (Gallais *et al.*, 2007). Mapping populations such as F_2, backcross, recombinant inbred lines (RILs) *etc.*, can be used and with precise

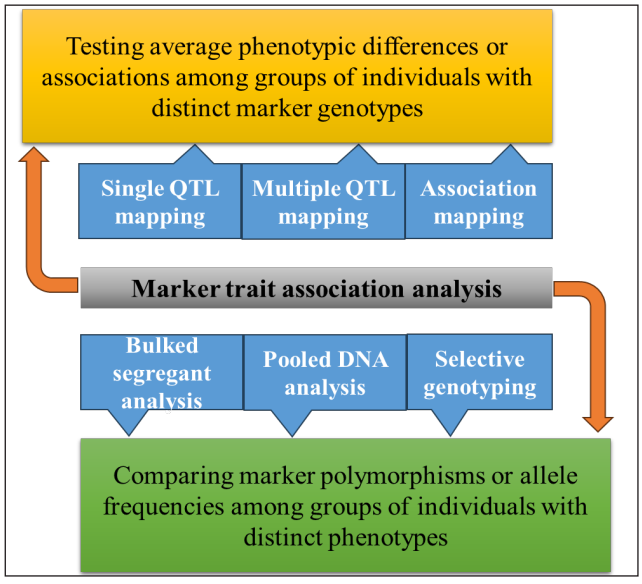

Figure 1: Key approaches for identifying marker-trait associations

Article History

RECEIVED on 25th December 2024 RECEIVED in revised form 03rd February 2025 ACCEPTED in final form 04th February 2025

phenotyping, 30 to 50 individuals exhibiting extremely high and low phenotypic values are selectively genotyped. Based on the significant difference in marker allele frequencies between the extreme groups, markers linked to the trait are identified.

Principle Underlying Selective Genotyping

Genotyping individuals exhibiting extreme phenotypes for marker-trait linkage establishment is based on the principle of genetic hitchhiking. In genetic hitchhiking, when a marker is linked to a trait, the respective marker alleles associated with the extreme phenotypic counterparts are expected to be co-selected during the formation of groups comprising individuals with extreme phenotypic values. For example, if marker A (with two alleles A_1 and A_2) is linked to flowering time and the mapping population is homozygous, such as RILs, in the extreme early-flowering group, most of the individuals will possess the A_1A_1 marker genotype, while in the extreme late-flowering group most of the individuals will possess the A_2A_2 marker genotype. A t-test comparing marker allele frequencies between the extreme phenotypic groups can determine whether the difference is significant or due to chance. For unlinked markers, the marker alleles are expected to be equally frequent between the early- and late-flowering groups (*i.e.*, no statistically significant difference in marker allele frequencies is observed).

A hypothetical example is provided in figure 2 to illustrate the principle underlying selective genotyping. To test the significant difference in marker allele frequencies between the high and low phenotypic value groups, each consisting of eight individuals, a two-sample, two-tail t-test is applied. The calculated t value is found to be 2.77, while the tabulated t value is 2.145 at 5% level of significance and 14 degrees of freedom ($n_1 + n_2 - 2 = 8 + 8 - 2 = 14$, where n_1 and n_2 represent the number of individuals in the high and low phenotypic value groups respectively). Since the calculated t-value is greater than the tabulated t-value, the alternate hypothesis is accepted and there is a significant difference in marker allele frequency between the high and low phenotypic value groups. Therefore, the presence of linkage between marker A and the trait for which the extreme groups were constituted can be deciphered.

Null hypothesis (H_0): $f(A_1)_{high}$ is equal to $f(A_1)_{low}$

Alternate hypothesis is (H_1): $f(A_1)_{high}$ is not equal to $f(A_1)_{low}$

High phenotypic value group

A_1	------	------	------	------	------	------	------	
A_2							------	------

$f(A_1)_{high} = (2\times6 + 1\times1)/(2\times8) = 0.81$

Low phenotypic value group

A_1	------	------				------		
A_2		------	------	------	------	------	------	------

$f(A_1)_{low} = (2\times1+1\times2)/(2\times8) = 0.25$

$$t_{cal} = \frac{|f(A_1)_{high} - f(A_1)_{low}|}{\sqrt{\frac{f(A_1)_{high}(1 - f(A_1)_{low}) + f(A_1)_{low}(1 - f(A_1)_{high})}{2 \times n}}}$$

$$\frac{|0.81 - 0.25|}{\sqrt{\frac{0.81(1 - 0.25) + 0.25(1 - 0.81)}{2 \times 8}}} = 2.77$$

Since t_{cal} (2.77) > t_{tab} (2.145) value at 5% level of significance, H_0 is rejected and H_1 is accepted.

Figure 2: Illustration of linkage detection between marker and trait using the selective genotyping method

Types of Selective Genotyping

Selective genotyping is categorized into bidirectional and unidirectional approaches (Navabi *et al.*, 2009).

1. Bidirectional Selective Genotyping

This method includes selecting 30 to 50 individuals for genotyping from each extreme tail of the population. It is suitable when both tails survive without significant mortality. In this case, whether any significant difference exists in allelic frequencies between the two tails is tested. A two-sample, two-tailed t-test is conducted to establish marker-trait associations. The example illustrated in figure 2 demonstrates conducting a t-test for bidirectional selective genotyping.

2. Unidirectional Selective Genotyping

In cases where high selection pressure or lethal alleles result in the mortality of one tail, a two-sample t-test is not feasible. This approach is useful in tagging markers linked to diseases resistance such as resistance to viral infections or wilts. Marker allele frequency obtained from the genotyped individuals of surviving tail is compared with an allele frequency of 0.5 using one-sample t-test. However, the QTL detection power of unidirectional selective genotyping is relatively lowers than bidirectional selective genotyping.

Factors Influencing QTL Detection Power

The QTL detection power of selective genotyping is affected by several factors such as size of extreme bulks, mapping population size, genome wide marker density and presence of linked QTLs. Thus, these parameters are important considerations before implementing selective genotyping (Sun *et al.*, 2010).

1. Extreme Bulk Size: Increasing the size of extreme bulks increases QTL detection power, but beyond an optimal size, the power either does not increase significantly or diminishes.

2. Population Size: Larger mapping population increases QTL detection power by reducing false positives in extreme bulks and thus increases statistical robustness.

3. Marker Density: Higher marker density improves QTL detection. However, the effect of phenotypic variance explained is higher than marker density for QTL detection.

4. Linkage: The power of selective genotyping to distinguish between two linked QTLs is greater in the repulsion phase than in the coupling phase. It has been found that inclusive composite interval mapping has higher resolution for detecting closely linked QTLs compared to selective genotyping.

Advantages and Limitations of Selective Genotyping

A substantial reduction in time and cost is feasible by implementing selective genotyping compared to

conventional QTL mapping. For example, genotyping only 30 individuals from each extreme phenotypic tail of a mapping population consisting of 1,000 individuals reduce the cost to 6% compared to genotyping the entire population. In practical breeding population, selection is practiced and QTL mapping is not followed. Selective genotyping has been suggested for utilization in QTL detection in breeding population as well (Sun *et al.*, 2010).

However, selective genotyping has several limitations compared to. Using a small number of markers for genome-wide coverage may not be feasible for easy QTL identification. Selecting contrasting individuals from small populations reduces QTL detection power, favouring the identification of large-effect QTLs. Additionally, gel-based genotyping systems may inaccurately quantify allele frequencies, particularly for rare alleles, reducing the detection accuracy.

Conclusion

Selective genotyping is a powerful and cost-effective method for QTL detection, offering significant advantages by avoiding genotyping the entire mapping populations. Its ability to identify makers linked to traits of interest with minimal genotyping effort makes it particularly valuable in breeding programs. Selective genotyping has been applied for QTL mapping for various traits, including yield-related traits in crops such as wheat (Yang *et al.*, 2020). Despite its limitations, such as reduced detection power in small populations and challenges with rare allele detection, selective genotyping can be a choice for breeders aiming to combine simultaneous breeding and QTL identification. Moreover, next-generation sequencing platforms have enabled extreme pool-based QTL identification approaches more feasible and techniques such as QTL-seq and MutMap, have gained significant recognition and are increasingly applied in plant breeding.

References

Gallais, A., Moreau, L., Charcosset, A., 2007. Detection of marker-QTL associations by studying change in marker frequencies with selection. *Theoretical and Applied Genetics* 114, 669-681. DOI: https://doi.org/10.1007/s00122-006-0467-z.

Navabi, A., Mather, D.E., Bernier, J., Spaner, D.M., Atlin, G.N., 2009. QTL detection with bidirectional and unidirectional selective genotyping: Marker-based and trait-based analyses. *Theoretical and Applied Genetics* 118, 347-358. DOI: https://doi.org/10.1007/s00122-008-0904-2.

Sun, Y., Wang, J., Crouch, J.H., Xu, Y., 2010. Efficiency of selective genotyping for genetic analysis of complex traits and potential applications in crop improvement. *Molecular Breeding* 26, 493-511. DOI: https://doi.org/10.1007/s11032-010-9390-8.

Yang, L., Zhao, D., Meng, Z., Xu, K., Yan, J., Xia, X., Zhang, Y., 2020. QTL mapping for grain yield-related traits in bread wheat via SNP-based selective genotyping. *Theoretical and Applied Genetics* 133, 857-872. DOI: https://doi.org/10.1007/s00122-019-03511-0.