JUST AGRICULTURE
multidisciplinary e-Newsletter

QTL Seq: In Delineation of Loci Governing Flowering Time

Shivaprasad K.M., ShashidharB. R.*,Harisha R. and Pavan Kumar Naik N.
Division of Genetics, ICAR – Indian Agricultural Research Institute New Delhi 110012

ARTICLE ID: 61

Abstract

Quantitative trait loci (QTL) mapping is a genome-wide inference for the relationship between genotypic DNA at various locations and phenotype for a set of quantitative traits in terms of the number, effects, positions, and interaction of QTL's. This localization helps for the ultimate identification of Genetic regulatory network (GRN) with understanding of mechanisms underneath. Traditionally, QTL mapping requires mapping population, genome-wise distributed polymorphic markers for linkage analysis and is time-consuming and labour-intensive. Recently a rapid approach for identification of such plant QTLs was developed which uses whole-genome resequencing (WGRS) of DNAs from two bulks of extreme trait values in opposite direction for a given phenotype in a segregating progeny. This is popularly known as QTL-seq which is a modification of BSA(bulked-segregant analysis).To tackle the food needs of exponentially growing population there is a great demand for improving farm produce without increasing agricultural land. Targeted Breeding for early maturing crops can drastically reduce crop maturation time, input resources and provide opportunity for additional cropping season. But the regulatory network for flowering time is complex and is responsive to multiple environmental cues, including photoperiod, ambient temperature, vernalization, and nutrient status. The transition from growth phase to the reproductive phase is tightly controlled by multiple genetic and physiological signals. Use of QTL seq can rapidly locate flowering time related genes and aid in architecting early maturing crops.

Introduction

The world's population is growing at rapid pace and has crossed 7 billion, while the land suitable for agriculture is decreasing because of several reasons such as rapid climate change, urbanization etc. Therefore, there is a high demand for efficient crop improvement to increase harvest index without further expanding farmland and damaging the environment. In

JUST AGRICULTURE
multidisciplinary e-Newsletter

plants, multiple genes each with a relatively differing effect control the majority of ergonomically important characters and are defined as quantitative trait loci (QTLs) (Falconer and Mackay, 1996). Identification of QTLs on the chromosomes is an important task in crop breeding. Once established a QTL,DNA markers linked to such regions can be identified and can be efficiently introduced into an elite cultivar to increase its positive traits as directed by man. Recurrent backcrossing of elite parent to the donor plant, followed by selection of progeny inheriting the desirable QTL a process known as marker-assisted selection can be used for the purpose. It reduces the effort and time needed for selection and evaluation of progeny and introduce required character without linkage drag by selecting for recombination. Traditionally QTL mapping has been widely adopted to locate QTL sby linkage analysis of progeny derived by selective cross between parents having highly contrasting phenotypes for the trait of interest. It requires DNA markers capable of discriminating parental genomes situated throughout the genomes. This requirement of developing mapping population by crosses involving genetically distantly related cultivars entails those parents may be differing in many QTLs controlling the phenotype and complicating isolation of individual loci. On the other hand, when closely related parents are used, identification of sufficient polymorphic markers becomes a limiting step.

History of QTL seq

Michelmore *et al*. in 1991 reported an elegant approach to identify markers tightly associated to the causal gene for a given phenotype. This method termed asbulked- segregant analysis (BSA) is used on scoring F2 progeny following a cross between contrasting parental lines showing contrasting phenotypes. Two bulked DNA for each extreme phenotype with at least 10 samples each are generated and polymorphic markers between parents are scored to identify a marker which retains polymorphism. Such marker can be validated to know its association with trait under study. Later, BSA was applied to identify QTLs, which is called ‘selective DNA pooling’(Mansur *et al*. 1993). However, usage and availability of DNA markers was the important factor limiting its effectiveness. Current development of whole genome sequencing (WGS) has accelerated the analysis of QTLs. In yeast, a model fungus with a relatively small genome size (12.5 Mb) was used by Ehrenreich *et al.* (2010) to make cross between two diploid yeast strains and obtained a large number of haploid progenies. Then BSA was applied to select two populations with extreme phenotypes, and genotyped

JUST AGRICULTURE
multidisciplinary e-Newsletter

with microarray (SNP) and WGS, which successfully identified the QTLs involved in resistance to several chemical compounds. The approach is called XQTL, since extremely large numbers of progenies were used in individualbulks. Similar applications of WGS to identify QTLs in rice with much larger genome sizes were reported by Takagi *et al*. (2013). They applied QTL-seq to rice RILs (recombinant inbred lines) and F2 populations and successfully identified QTLs for important agronomic traits, such as partial resistance to the fungal rice blast disease and seedling vigour.

Flowering pathway

Flowering time in *A. thaliana* is a complex trait that response to multiple internal and environmental factors. Externally it is regulated by photoperiod, vernalization and temperature while internal factors such as age and gibberellic acid affect flowering. Floral integrator genes receive signals from all these factors, during positive signals FLT is activated, commonly known as florigen. The translated proteins moves from leaves to flowers and dimerizes with FD and activate floral meristem identity genes such as LEAFY (LFY) and APATALA1 (AP1). This activation results in conversion of vegetative meristems into floral meristem further activate floral organ identity genes such as AP1, AP2, AP3, PISTILLATA and AGAMOUS, which leads to the development of flower. In photoperiod and vernalization pathway, normally in the leaves the floral CONSTANT (CO) protein regulates the activity of FLT, which is affected by combined action of circadian clock and photoperiod. During dark conditions CO protein is degraded, but when the plants exposed to light FKF1 protein absorbs light and bind to the CO and prevents its degradation. This makes sure CO protein accumulation in long day conditions, which activate and mobilises FLT to meristem from leaves and cause flowering. Whereas, Vernalization pathway is mainly regulated by FLC gene, FLC inhibits flowering by binding regulatory elements of key flowering genes, FLC is repressed by cold. So vernalizations eliminate FLC leading to flowering.

QTL Seq in mapping of flowering time

Flowering time is an important trait and the target of selection in plant breeding. Genetic control of flowering time is complex and quantitative in nature. Many studies have detected quantitative trait loci (QTL) related to flowering time or earliness in various crops. Proper timing of flowering has an important adaptive value for flowering plants. The transition from the vegetative phase to the reproductive phase is tightly controlled by multiple

JUST AGRICULTURE
multidisciplinary e-Newsletter

physiological signals and genetic pathways. Flowering time genes represents one of the best studied functional genetic networks in plants, as geneticists have identified 60 genes that regulate flowering time (Baurle and Dean 2006). Early flowering being an important agronomic trait in cucumber, Lu *et al* (2014) in their study, identified QTL for early flowering. Ef1.1 an important gene on chromosome 1 was located through QTL-seq, which was confirmed by microsatellite marker-based classical QTL mapping in the F2 population in cucumber. Joint QTL-seq and traditional QTL analysis delimited Ef1.1 to an 890 kb genomic region.Csa1G651710 a candidate gene was identified in this region, which is a homolog of the FLOWERING LOCUS T (FT), the main flowering switch gene in Arabidopsis. Validating through expression level of Csa1G651710 revealed significant higher expression in early flowering genotypes, a possible candidate gene for early flowering in the cucumber line Muromskij. Employing QTL-seq strategy, Shu *et al* (2018) identified a major genomic region harbouring a robust flowering time QTL using an F2 mapping population, designated Ef2.1 on broccoli chromosome 2. Combined QTL-seq and classical QTL analysis narrowed down Ef1.1 to a 228-kb genomic region containing 29 genes. A broccoli gene, Bol024659, was identified in this region, which is a homolog of GRF6, a major gene regulating flowering in Arabidopsis, and was designated BolGRF6. The identified candidate genomic regions and genes may be useful for molecular breeding to improve broccoli and cabbage flowering times. Ruangrak *et al* (2017) identified a candidate QTL of an early flowering trait in tomato using QTL-seq. Using traditional QTL analysis, the location of one QTL was confirmed in the physical region between 23.5 and 25.3 Mb, which corresponded to the region identified using QTL-seq, and was referred to as EF1 (Solyc01g017060). Validation by qRT-PCR analysis showed that EF1 was the highly expressed gene among the candidate genes in early flowering parents.Use of such techniques to map flowering time related genes, will help in better understanding the mechanism in crop plants, further assisting in development of superior early flowering cultivars.

QTL Seq

QTL Seq is the technique for QTL identification using whole-genome resequencing (WGRS) of two DNA bulks of progeny (each with 20–50 individuals) showing extreme phenotypic values by next-generation sequencing (NGS) technology.

Steps involved in QTL Seq

JUST AGRICULTURE
multidisciplinary e-Newsletter

1. Development of mapping population: Develop mapping population by crossing individuals with contrasting phenotype for target trait. Most commonly used mapping population include F2 or RILs.
2. Phenotyping of the mapping population
3. DNA from 10 to 20 individuals from each extreme are bulked to generate 'Highest' bulk and 'Lowest' bulk
4. WGRS of each bulked DNAs separately with genome coverage of at least 6x.
5. Alignment of short reads to the reference sequence of cultivar using BWA software
6. Calculating the SNP-index by SNP calling using SAM tools: SNP-index is the proportion of reads harbouring the SNP that are different from the reference sequence.
7. Plot the SNP-index graph using a sliding window analysis with 1 Mb window size and 10 kb increment.
8. Calculate the Δ(SNP-index): Δ(SNP-index) =SNP-index of (H-bulk –L-bulk)
9. Plot the graph of Δ(SNP-index)
10. The QTL can be identified as peaks or valleys of the SNP-index plot or by identifying the genomic region which has Δ(SNP-index) value of 1 / -1. Whereas genomic region which do not contain QTL has Δ(SNP-index) value of 0.

QTL seq and Flowering time in Crop improvement

QTL seq allows rapid identification of QTLs, as DNA marker development and genotyping is not required, which is the most time-consuming and costly procedure needed for the conventional QTL analysis. While cultivars with early flowering and maturity escape terminal drought, reduce the amount water, pesticides, cost of cultivation and gives an opportunity for flexible multiple cropping system.

Conclusion

In view of the recent rapid development in sequencing technology, we foresee those methods that make use of whole-genome sequencing-based techniques, including QTL-seq, Mut Map (Abe *et al*,.2012a) and others will dramatically accelerate crop improvement in a cost-effective manner. These technologies that take full advantages of the rapidly declining cost of genome sequencing are expected to significantly contribute to the on-going efforts aimed at addressing the world food security problem by reducing breeding time. Since flowering time is a quantitative trait governed by polygenes, based on the SNP-index and

JUST AGRICULTURE
multidisciplinary e-Newsletter

Δ(SNP index) charts obtained from QTL-seq will be helpful to map QTLs responsible for flowering time variation in crop plants.

References

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M. and Innan, H., 2012. Genome sequencing reveals agronomically important loci in rice using MutMap. *Nature biotechnology*, *30*(2), pp.174-178.

Bäurle, I. and Dean, C., 2006. The timing of developmental transitions in plants. *Cell*, *125*(4), pp.655-664.

Ehrenreich, I.M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J.A., Gresham, D., Caudy, A.A. and Kruglyak, L., 2010. Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature*, *464*(7291), pp.1039-1042.

Falconer, D.S. and Mackay, T.F.C., 1996. Introduction to Quantitative Genetics, 4th edn. London: Prentic.

Lu, H., Lin, T., Klein, J., Wang, S., Qi, J., Zhou, Q., Sun, J., Zhang, Z., Weng, Y. and Huang, S., 2014. QTL-seq identifies an early flowering QTL located near Flowering Locus T in cucumber. *Theoretical and applied genetics*, *127*(7), pp.1491-1499.

Mansur, L.M., Orf, J. and Lark, K.G., 1993. Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbreds of soybean (Glycine max L. Merr.). *Theoretical and Applied Genetics*, *86*(8), pp.914-918.

Michelmore, R.W., Paran, I. and Kesseli, R., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the national academy of sciences*, *88*(21), pp.9828-9832.

Ruangrak, E., Su, X., Huang, Z., Wang, X., Guo, Y., Du, Y. and Gao, J., 2018. Fine mapping of a major QTL controlling early flowering in tomato using QTL-seq. *Canadian Journal of Plant Science*, *98*(3), pp.672-682.

Shu, J., Liu, Y., Zhang, L., Li, Z., Fang, Z., Yang, L., Zhuang, M., Zhang, Y. and Lv, H., 2018. QTL-seq for rapid identification of candidate genes for flowering time in broccoli× cabbage. *Theoretical and applied genetics*, *131*(4), pp.917-928.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S. and Innan, H., 2013. QTL- seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *The Plant Journal*, *74*(1), pp.174-183.