Available online at www.sciencedirect.com

ScienceDirect

Genetic mapping of quantitative trait loci in crops

*Yang Xu, Pengcheng Li, Zefeng Yang, Chenwu Xu**

Jiangsu Provincial Key Laboratory of Crop Genetics and Physiology/Co-Innovation Center for Modern Production Technology of Grain Crops, Key Laboratory of Plant Functional Genomics of Ministry of Education, Yangzhou University, Yangzhou 225009, China

ARTICLE INFO

Article history:
Received 12 May 2016
Received in revised form
20 June 2016
Accepted 11 July 2016
Available online 21 July 2016

Keywords:
Family-based mapping
Natural population-based mapping
Mixed linear model
MAGIC population
Meta-analysis
Genotyping by sequencing

ABSTRACT

Dissecting the genetic architecture of complex traits is an ongoing challenge for geneticists. Two complementary approaches for genetic mapping, linkage mapping and association mapping have led to successful dissection of complex traits in many crop species. Both of these methods detect quantitative trait loci (QTL) by identifying marker–trait associations, and the only fundamental difference between them is that between mapping populations, which directly determine mapping resolution and power. Based on this difference, we first summarize in this review the advances and limitations of family-based mapping and natural population-based mapping instead of linkage mapping and association mapping. We then describe statistical methods used for improving detection power and computational speed and outline emerging areas such as large-scale meta-analysis for genetic mapping in crops. In the era of next-generation sequencing, there has arisen an urgent need for proper population design, advanced statistical strategies, and precision phenotyping to fully exploit high-throughput genotyping.

1. Introduction

The objective of genetic mapping is to identify QTL responsible for natural phenotypic variation. Two strategies have been widely applied to genetic mapping in plants: (1) linkage mapping and (2) association or linkage disequilibrium (LD) mapping. Linkage mapping, a conventional mapping method, depends upon genetic recombination during the construction of mapping populations. Over the past two decades, linkage mapping has been commonly used in various plant species, and many QTL have been cloned or tagged [1]. However, linkage mapping has the disadvantages of relatively low mapping resolution, low allele richness, and low speed.

Association mapping, as a complement to linkage mapping, takes advantage of historic recombination events accumulated over hundreds of generations, thus providing higher resolution and greater allele numbers [2]. Since human diseases were successfully dissected, association mapping has been applied to crops [3]. Following its introduction to crops [4], association mapping has attracted increased attention in genetic studies. Owing to the dramatic reduction in costs of sequence technologies, association mapping has been conducted in plants from the model plant *Arabidopsis thaliana* [5] to many major crops, such as rice [6], maize [7], wheat [8], soybean [9], barley [10], sorghum [11], potato [12], and tomato [13].

The key distinction between association and linkage mapping lies in whether recombination events occur in populations or families. However, both of these methods share a consistent strategy for identifying molecular markers that are linked to QTL. As we step into the era of complete genome sequencing, the

* *Corresponding author.*
E-mail address: cwxu@yzu.edu.cn (C. Xu).
Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

http://dx.doi.org/10.1016/j.cj.2016.06.003

difference between the two methods will disappear [14]. Genetic mapping can be generally classified into family-based mapping when mapping is performed in progenies of biparental or multiparent crosses and natural population-based mapping when mapping is conducted in natural populations in which relationships are unknown. In this review, we describe the family-based mapping and natural population-based mapping of complex traits, highlight the statistical methods used for genetic mapping, and outline the developmental trends and perspectives of genetic mapping in crop genetics.

2. Family-based mapping

2.1. Biparental populations

The first and most important step in family-based mapping is constructing experimental populations, which may be biparental populations such as F_2, backcrosses (BC), doubled haploids (DH), recombinant inbred lines (RIL), and near-isogenic lines (NIL). These commonly used populations with their strengths and weaknesses are described in Table 1. The general process of biparental mapping includes: (1) collection of parental strains that differ for traits of interest, (2) selection of molecular markers such as RFLP, SSR and SNP that distinguish between the two parents, (3) development of a mapping population, (4) genotyping and phenotyping of the mapping population; and (5) detection of QTL using a suitable statistical method. The power of QTL detection is affected by QTL effects, allele frequencies, and the type and size of the mapping population. Biparental mapping has proven to be useful in crop breeding [15]. The main limitation of a biparental population is that only a few recombination events occur during the development of the population, allowing the localization of QTL to 10–20 cM intervals. Additionally, detection of QTL in biparental populations depends on the phenotypic diversity of the two parents, which may account for only a small part of the genetic variation in the species.

Table 1 – Commonly used biparental populations with their strengths and weaknesses.

Population	Strengths	Weaknesses
F_2	Rapid construction, estimation of both additive and dominant effects	Lower power, limited recombination, temporary nature
BC	Utility for introgressing specific genes	Impossibility of estimation of dominant effects, time requirement, temporary nature
DH	Rapid construction, immortality, easy replication	Limited recombination, expense, impossibility of estimation of dominant effects
RIL	Abundance of recombination, immortality, easy replication	Impossibility of estimation of dominant effects, time requirement

2.2. Multiparent mapping populations

Multiparent mapping populations have been constructed to overcome the limitations of biparental populations. The genetic diversity of multiple parents leads to a population with large phenotypic diversity, making it suitable for high-resolution QTL mapping. Increasing in popularity are two experimental designs of multiparent populations that include nested association mapping (NAM) and multiparent advanced generation intercrosses (MAGIC). Recently published multiparent mapping studies in crops are summarized in Table 2.

NAM is an excellent multiparent population design suggested by Yu et al. [16] for dissecting the genetic architecture of maize flowering time. A NAM population was created by crossing 25 diverse inbred maize lines to the B73 inbred, chosen as a reference line, resulting in 5000 RILs from 25 families, with 200 RILs per family. As a combination of several high-resolution biparental populations in one large population, the NAM population affords very high resolution and power for detecting QTL. In maize, a NAM population has been used for large-scale genetic mapping for several important traits including leaf architecture and disease resistance [17–19]. The use of MAGIC populations was first proposed for QTL mapping in mouse by Threadgill et al. [20]. In crops, Kover et al. [21] first developed a MAGIC population in *A. thaliana* that consisted of 527 lines derived by intermating a heterogeneous panel of 19 founders. MAGIC populations have been used for identification of QTL for hectoliter weight and plant height in wheat [22]. MAGIC populations including several indica and japonica rice parents have been developed for QTL mapping and varietal development in rice [23]. Compared with other multiparent populations, MAGIC populations involve intermating multiple inbred founders for multiple generations prior to the construction of inbred lines, considerably improving the precision of QTL detection. Undoubtedly, MAGIC populations offer great opportunities for dissecting complex traits and improving breeding populations. Statistical approaches for QTL mapping in MAGIC populations have become available, some of them based on the general linear model (GLM) used in biparental populations [24].

3. Natural population-based mapping

With the advantages of high resolution, high allelic richness, and absence of need of the tedious development of a mapping population, natural population-based mapping has become a powerful tool for detection of natural variation underlying complex traits in more than a dozen crops since 2001. The main steps in natural population-based mapping are depicted in Fig. 1. They consist of first, collection of a sample population including elite cultivars, landraces, wild relatives, and exotic accessions; second, phenotyping traits, estimating broad-sense heritability of traits of interest and determining the genotypes of the population entries, either for candidate genes or genome-wide; third, quantification of the LD extent of the selected population; fourth, identification of the influence of population structure and kinship; and fifth, testing the associations between genotypes and phenotypes using appropriate statistical approaches. Subsequent experimental validations such as mutagenesis and gene