Field Crops Research 134 (2012) 185–192

Contents lists available at SciVerse ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Bulk segregant analysis: "An effective approach for mapping consistent-effect drought grain yield QTLs in rice"

Prashant Vikram[a], B.P. Mallikarjuna Swamy[a], Shalabh Dixit[a], Helaluddin Ahmed[a,1], M.T. Sta Cruz[a], Alok K. Singh[b], Guoyou Ye[c], Arvind Kumar[a,*]

[a] Plant Breeding Genetics & Biotechnology Division, International Rice Research Institute, DAPO Box 7777, Metro Manila, Philippines
[b] Department of Genetics and Plant Breeding, V.B.S. Purvanchal University, Jaunpur, India
[c] Crop Research Informatics Laboratory, International Rice Research Institute, DAPO Box 7777, Metro Manila, Philippines

ARTICLE INFO

Article history:
Received 26 February 2012
Received in revised form 30 May 2012
Accepted 30 May 2012

Keywords:
QTLs
Quantitative trait loci
Selective genotyping
Bulk segregant analysis
Grain yield under drought

ABSTRACT

Mapping QTLs for grain yield under drought in rice involves phenotyping and genotyping of large mapping populations. The huge cost incurred in genotyping could be considered as a bottleneck in the process. Whole population genotyping (WPG), selective genotyping (SG), and bulk segregant analysis (BSA) approaches were employed for the identification of major grain-yield QTLs under drought in rice in the past few years. The efficiency of different QTL mapping approaches in identifying major-effect grain-yield QTLs under drought in rice was compared using phenotypic and genotypic data of two recombinant inbred line populations, Basmati334/Swarna and N22/MTU1010. All three genotyping approaches were efficient in identifying consistent-effect QTLs with an additive effect of 10% or more. Comparative analysis revealed that SG and BSA required 63.5% and 92.1% fewer data points, respectively, than WPG in the N22/MTU1010 $F_{3:4}$ mapping population. The BSA approach successfully detected consistent-effect drought grain-yield QTLs $qDTY_{1.1}$ and $qDTY_{8.1}$ detected by WPG and SG. Unlike SG, BSA did not lead to an upward estimation of the additive effect and phenotypic variance. The results clearly demonstrate that BSA is the most efficient and effort saving genotyping approach for identifying major grain yield QTLs under drought.

1. Introduction

Advances in molecular marker technology have enabled fast-track improvement of crop plants in recent years. Marker-based approaches, including marker-assisted backcrossing and QTL pyramiding, have been applied in cereals for improving the tolerance of biotic and abiotic stresses (Collard and Mackill, 2008; Ye and Smith, 2010). Drought is an important abiotic stress causing huge losses in rice yields. Progress in breeding rice for drought tolerance could be made more efficient by applying marker-assisted breeding (Bernier et al., 2007). To unveil the genetic basis of a complex trait such as grain yield under drought, it is a prerequisite to genotype and phenotype mapping populations consisting of a large number of individuals. The costs associated with large-scale phenotyping and genotyping are a bottleneck to conducting a study on several populations at any particular time in identifying QTLs with consistent effects across multiple environments and genetic backgrounds. With little alternative available to phenotype in different seasons/environments, the costs involved with genotyping could be effectively reduced by successfully applying different genotyping approaches.

In a model cereal crop such as rice (*Oryza sativa*) with a genome size of 389 Mb, it is a common practice to use 200–400 lines (recombinant inbred lines, backcross inbred lines, doubled haploids) as a mapping population. A dense map covering all 12 chromosomes with an average genetic distance of 10–15 cM has been developed to identify the precise QTLs. In most of the QTL mapping experiments, the whole genome is scanned (Gomez et al., 2010; Lanceras et al., 2004; Xu et al., 2005) to ultimately find only a few significant markers showing association with the trait under study. To reduce the extra burden and costs associated with genotyping of large mapping populations, alternative strategies have been proposed. A trait-based genotypic analysis called "selective genotyping" (SG) was suggested by Lebowitz et al. (1987), in which progenies are categorized into distinct classes based on the trait

Abbreviations: BSA, bulk segregant analysis; MAS, marker-assisted selection; QTLs, quantitative trait loci; SG, selective genotyping; WPG, whole population genotyping.

* Corresponding author. Tel.: +63 2 580 5600x2586; fax: +63 2 580 5699.
E-mail addresses: p.vikram@irri.org (P. Vikram), m.swamy@irri.org (B.P.M. Swamy), s.dixit@cgiar.org (S. Dixit), helaluddinahmed@hotmail.com (H. Ahmed), m.stacruz@cgiar.org (M.T.S. Cruz), alok6619@gmail.com (A.K. Singh), G.Ye@cgiar.org (G. Ye), akumar@cgiar.org (A. Kumar).
[1] Present address: Plant Breeding Division, Bangladesh Rice Research Institute, Joydebpur, Gazipur 1701, Bangladesh.

0378-4290/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.fcr.2012.05.012

values and marker allele frequencies are compared between the classes. In this approach, a subset of the population is used for genotyping instead of the whole population (Lander and Botstein, 1989). This approach was used for mapping major-effect drought grain-yield QTL *qDTY*$_{12.1}$ using only 169 (38.7%) lines from a population size of 436 (Bernier et al., 2007); later, this QTL was reconfirmed with only 4.5% of the lines from the whole population (Navabi et al., 2009). Another time and effort saving approach is bulk segregant analysis (BSA) (Michelmoore et al., 1991). In BSA, DNA of progenies corresponding to the phenotypic extremes is extracted and pooled. Therefore, only two pools of extreme lines along with the parents are genotyped for the identification of markers linked with the trait of interest. BSA was first applied to the identification of markers linked with disease resistance. Initially, it was applied to the diseases in which resistance was mostly governed by major genes, usually qualitative in nature (Michelmoore et al., 1991). Recently, BSA has been applied for quantitative traits also such as QTLs for heat tolerance in rice (Zhang et al., 2009), salt tolerance in Egyptian cotton (El-Kadi et al., 2006), and drought tolerance in wheat and maize (Altinkut and Gozukirmizi, 2003; Quarrie et al., 1999; Kanagaraj et al., 2010; Venuprasad et al., 2009; Vikram et al., 2011), as well as QTLs for grain yield under drought in maize (Quarrie et al., 1999). BSA has been used for the identification of QTLs associated with high grain yield under drought in rice (Venuprasad et al., 2009, 2011; Vikram et al., 2011).

Whole population genotyping involves markers from all the 12 chromosomes so that they represent whole genome and genetic background is taken in to account while QTL analysis. Contrarily, BSA and SG (selective genotyping) approaches may not consider all recombination options during QTL identification due to the unavailability of additional marker data on the same chromosome and other chromosomes. However, these approaches have been proven to be powerful enough to identify major QTLs that are worthy for MAS (Bernier et al., 2007; Venuprasad et al., 2009; Vikram et al., 2011). BSA and SG might fail to detect QTLs with smaller effects because only the extreme or tail lines are used for analysis. Also, interactions between different loci are least likely to be detected through these approaches. QTL analysis through BSA or SG approaches does not takes in to account the background so that they might tilt upward the magnitude of phenotypic variance, LOD (logarithm of odds) score, and additive effects.

The study was undertaken to perform a comparative analysis of SG and BSA in identifying major-effect QTLs for grain yield under drought and compare the fluctuations in phenotypic variance, additive effect, LOD, and level of significance values obtained in BSA and SG compared with WPG (whole population genotyping). WPG, SG, and BSA approaches were compared in a population derived from a traditional variety (Basmati334) and a popular lowland rice variety (Swarna). Comparative analysis of these approaches might be biased if only one population is taken into account. To validate the results and perform a comprehensive analysis, phenotypic and genotypic data of a rice population applied to QTL mapping in earlier studies were used (Vikram et al., 2011).

2. Materials and methods

The study was conducted at the International Rice Research Institute (IRRI), Los Baños, Laguna, Philippines. An $F_{3:4}$ Basmati 334/Swarna population was phenotyped for grain yield under drought and genotyped through WPG, SG, and BSA. Basmati334 is a drought-tolerant local landrace whereas Swarna is a popular lowland rice cultivar in India (Sivaranjani et al., 2010; Verulkar et al., 2010). The phenotypic and genotypic data of the N22/MTU1010 population were also used for comparison of WPG, SG, and BSA (Vikram et al., 2011). Basmati334/Swarna population was screened under drought stress in wet season 2008 (WS2008) and N22/MTU1010 in dry season of 2009 (DS2009). WS2008 experiments were shown on June 17, 2008 whereas, DS2010 experiments on December 9, 2009.

2.1. Phenotyping of Basmati334/Swarna population

The $F_{3:4}$ Basmati334/Swarna RIL population was phenotyped for grain yield under drought stress as well as irrigated/non-stress conditions (Supplementary Table 1). The numbers of lines used for drought screening were 204 during WS2008 and 367 during DS2010. The 204 lines used in WS2008 were a subset of 367 lines that were phenotyped in DS2010. Screening for grain yield under drought in the lowland rice ecosystem was carried out at IRRI using a standard phenotyping protocol (Kumar et al., 2008; Venuprasad et al., 2007). Data for grain yield ($g\,m^{-2}$ converted to $kg\,ha^{-1}$), days to 50% flowering (DTF), plant height (PH), biomass (BIO), and harvest index (HI) were recorded.

2.2. Genotyping of Basmati334/Swarna population

Leaf samples were collected from a whole F_4 plot of the non-stress experiment and bulked. DNA was extracted by the modified CTAB (cetyl tri-methyl ammonium bromide) method Murray and Thompson, 1980). DNA was quantified on 0.8% agarose gel to adjust the concentration to 25 ng μL^{-1}. Quantified DNA was subjected to PCR amplification with a 15-μL reaction mixture involving 50 ng DNA, 1× PCR buffer, 100 μM dNTPs, 250 μM primers, and 1 unit Taq polymerase enzyme. Eight percent non-denaturing PAGE was used for the resolution of PCR products (Sambrook et al., 1989). A parental polymorphism survey between Basmati334 and Swarna was carried out with 880 simple sequence repeat (SSR) markers (Temnykh et al., 2001; McCouch et al., 2002; IRGSP, 2005). Genotyping of the Basmati334/Swarna population of 367 lines was done by 71 polymorphic SSR markers spread throughout the 12 chromosomes of the rice genome. BSA was also done for QTL identification with only 10% of the lines from the population (5% with high yield and 5% with low yield under drought stress). Grain yield data from the stress trial of DS2010 were used for selecting lines to constitute BSA. DNA of all the selected lines was quantified and bulked in equal quantity to prepare high- and low-yield bulks. DNA of these two bulks was screened with 203 polymorphic SSR markers along with parents – Basmati334 and Swarna (Fig. 1). A similar procedure for BSA in the N22/MTU1010 population was carried out (Vikram et al., 2011). Polymorphic markers from all the 12 chromosomes were selected at equal distance so that they represent whole genome. SG was carried out with the same number of markers used in WPG. Both RILpulations-Basmati334/Swarna and N22/MTU1010 presented in the study were analyzed through whole genome and BSA both.

2.3. Statistical analysis

Statistical analysis was performed through SASv9.1.3 (SAS Institute Inc., 2004). The REML algorithm of PROC MIXED of SAS v9.1.3 was used for the determination of mean values of entries using a model in which lines were treated as a fixed effect and replications and blocks within replications as random. Inbred line mean-based broad-sense heritability (H) was computed as:

$$H = \frac{V_g}{V_g + V_e/r}$$

where V_g is genetic variance, V_e is error variance, and r is the number of replications.