

Review

Germplasm-regression-combined marker-trait association identification in plants

Chengjiang Ruan

Dalian Nationalities University, Institute of Bio-Resources and Environment, No.18 Liaohe Western Road or Dalian Economic and Technical Development Zone, Dalian City, Liaoning Province 116600, P.R. China.

E-mail: ruancj@yahoo.com.cn. Fax: +86-411-87618179.

Accepted 31 December, 2009

In the past 20 years, the major effort in plant breeding has changed from quantitative to molecular genetics with emphasis on quantitative trait loci (QTL) identification and marker assisted selection (MAS). However, results have been modest. This has been due to several factors including absence of tight linkage QTL, non-availability of mapping populations and lack of substantial time needed to develop such populations. To overcome these limitations and as an alternative to planned populations, molecular marker-trait associations have been identified by the combination between germplasm and the regression technique. In the present preview, we first surveyed the successful applications of germplasm-regression-combined (GRC) molecular marker-trait association identification in plants; secondly, we described how to do the GRC analysis and its differences from mapping QTL based on a linkage map reconstructed from the planned populations; thirdly, we considered the factors that affect the GRC association identification, including selections of optimal germplasm and molecular markers and testing of identification efficiency of markers associated with traits; and finally we discussed the future prospects of GRC marker-trait association analysis used in plant MAS/QTL breeding programs, especially in long-juvenile woody plants when no other genetic information such as linkage maps and Quantitative Trait Loci are available.

Key words: Association analysis, marker assisted selection (MAS), molecular marker, quantitative trait, germplasm-regression.

INTRODUCTION

The improvement of quantitative traits (that is productivity, disease resistance, abiotic stress tolerance and/ or quality) has long intrigued plant breeders. Many desirable varieties and lines have been selected and bred by conventional breeding programs, such as systematic selection, cross and mutation breeding. With a pedigree breeding program, the breeder will cross two parents and practice selection until advanced-generation lines with the best

phenotype for the quantitative trait under selection are identified. These lines will then be entered into a series of replicated trials to further evaluate the material with the goal of releasing the best lines as a cultivar. This conventional breeding program requires large inputs of labor, land and financial resources. For these reasons, plant breeders are motivated to identify promising lines as early as possible in the selection process.

In the past 20 years, the major effort in breeding has changed from traditional phenotypic-pedigree-based selection systems to molecular genetics with emphasis on quantitative trait loci (QTL) identification and marker assisted selection (MAS). MAS, which uses DNA markers to select optimal genotypes, is an excellent tool for selecting beneficial genetic traits that are difficult to measure, that exhibit low heritability and/or are expressed late in development (Ribaut and Hoisington, 1998; Davies et al., 2006; Wilde et al., 2007; Ender et al., 2008; Knoll and Ejeta, 2008), as well as for assessing the genetic potential

Abbreviations: QTL, quantitative trait loci; MAS, marker assisted selection; GRC, germplasm-regression-combined; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; ISSR, inter-simple sequence repeats; SAMPL, selective amplification of microsatellite polymorphic loci; SSR, simple sequence repeat; SCAR, sequence characterized amplified region; SNP, single nucleotide polymorphism; PCR, polymerase chain reaction.

of specific genotypes prior to phenotypic evaluation (Gebhardt et al., 2004). Molecular markers linked with QTL/major genes for traits of interest are being routinely developed in many crops and trees, using materials derived from planned crosses such as F_1 s (trees), F_2 s (progeny from F_1 selfing), BCs (back crossed progeny), BILs (F_2 selfed and back-crossed progeny), RILs (progeny from F_2 selfing) and NILs (F_2 back-crossed and selfed progeny). Some of these markers have been successfully applied in some plants breeding programs (Collard and Mackill, 2008) by Marker-assisted backcrossing (Helguera et al., 2003; Benchimol et al., 2005) and pyramiding (Ashikari et al., 2005; Zhang et al., 2006; Barloy et al., 2007) and some will be hopefully used in future MAS (Ren et al., 2005; Song et al., 2007; Hansen et al., 2008).

However, results of MAS/QTL have been modest (Kearsey and Farquhar, 1998; Collard and Mackill, 2008; Hospital, 2009), especially as there are still few reports on success stories in long-juvenile woody plants. This may be because (i) in linkage-based QTL analyses, non-availability of mapping populations (only few genotypes that are used as parents of mapping populations may not be effective in different backgrounds (Liao et al., 2001; Steele et al., 2006)), lack of substantial time needed to develop such, and planned populations (especially for long-juvenile woody plants) are major limitations in the identification of molecular markers for specific traits; (ii) the success of MAS/QTL largely depends on the extent of genetic linkage between markers and relevant QTL loci (Virk et al., 1996; Misztal, 2006), but the absence of tight linkage is in mostly linkage-based studies (Thomas, 2003); (iii) the existence of QTL \times environment interactions (Bouchez et al., 2002). In order to overcome these limitations and as an alternative to planned populations, molecular marker-trait association identifications have been conducted through the combination between the present germplasm and the regression technique (Wright and Mowers, 1994; Yonash et al., 2000; Chatterjee and Pradeep, 2003; Chatterjee and Mohandas, 2003; Pradeep et al., 2007; Srivastava et al., 2007) and increasingly adopted in many plants (Maureira-Butler et al., 2007). The germplasm-regression-combined (GRC) association studies not only allow mapping of genes/QTLs with higher level of confidence, but also allow detection of genes/QTLs, which will otherwise escape detection in linkage-based QTL studies based on the planned populations.

In the present preview, we first surveyed the successful applications of GRC molecular marker-trait association identification in plants and then describe how to conduct this analysis and the difference from mapping QTL based on a linkage map reconstructed from the planned population. We then considered the factors that affect GRC association identifications, including selection of optimal germplasm and molecular markers and testing of markers associated with desirable traits. In the end, we discussed

briefly the future prospects of GRC marker-trait association analysis used in plant MAS/QTL breeding programs, especially in long juvenile and long-lived woody plants when no other genetic information such as linkage maps and Quantitative Trait Loci are available.

EXEMPLIFICATIONS OF GRC MARKER-TRAIT ASSOCIATION IDENTIFICATION

Use of GRC analysis to identify the associations of molecular markers with desirable traits is increasingly being used in crops and woody plants with great success, such as Asia rice (Virk et al., 1996), wheat (Roy et al., 2006), tea (Mishra and Sen-Mandi, 2004), Alfalfa (Obert et al., 2000; Maureira-Bulter et al., 2007), mulberry (Vijayan et al., 2006; Kar et al., 2008), coconut (Shalini et al., 2007), birch (Wang, 2007; Wang et al., 2008; Xia et al., 2008), oat (Achleitner et al., 2008) and sea buckthorn (Ruan et al., 2009) (Table 1). These studies suggested that associations between DNA fragments and quantitative traits (QTLs) can be detected without the necessity of developing the planned mapping populations and reconstructing the linkage map.

Traits in these studies shown in Table 1 include yield and quality traits, disease and pest resistance, and drought tolerance. Molecular markers used in these studies involve random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), inter-simple sequence repeats (ISSR), selective amplification of microsatellite polymorphic loci (SAMPL) and simple sequence repeat (SSR). The mostly regressive methods used in the GRC association identifications are multiple regression analysis (MRA), also including single linear regression analysis and general and mixed linear model. Sampling accessions with different phenotypic traits shown in Table 1 are selected from (i) different genotypes of one species growing in different regions, (ii) different genotypes of multiple species and (iii) different individuals and clones of the same genotype of one species.

HOW TO DO GRC MARK-TRAIT ASSOCIATION IDENTIFICATION

The linkage-based QTL development is the selection of two parents that differ markedly in a particular quantitative character and then the determination of associations between markers and that character in F_2 or backcross progeny (Figure 1). In contrast, the GRC marker development is based on the phenotypic evaluation for traits of germplasm and associations between DNA fragments and traits can be detected without the construction of linkage-map based on the planned mapping populations (Figure 1). The time and cost used in the GRC analysis are significantly lower than that of the linkage-based QTL

Table 1. Molecular marker-trait association identifications by the combination between germplasm and regression technique in plants.

Plants	Traits	Materials	Markers	Regression	Markers associated with traits	References
Asian rice (<i>Oryza sativa</i>)	Culm length (CL) and number (CN), days to flowering (DF), grain width (GW), and panicle (PL) and leaf length (LL)	48 accessions from the South and Southeast Asia	RAPD	MRA	12 markers associated with CL, 12 for CN, 29 for DF, 32 for GW, 25 for PL and 16 for LL	Virk et al., 1996
Alfalfa (<i>Medicago sativa</i> L.)	Resistant and susceptible downy mildew	36 resistant plants from UC-143, and 36 susceptible plants from UC-123	AFLP	χ^2 analysis	Two markers associated with resistance and 2 for susceptibility	Obert et al., 2000
Tea (<i>Camellia</i>)	Drought tolerance	29 Darjeeling-grown clones	RAPD	MRA	OPAH02 associated with high activity of drought tolerance-specific isozymes at 99.9% confidence level	Mishra and Sen-Mandi, 2004
Bread wheat	14 agronomic traits including plant height (PH), days to flowering (DF), days to maturity (DM), tiller number (TN), flag leaf area (FLA), peduncle (PL) and spike length (SL), spikelets (SP) and florets (FL) and grain (GR) per spike, biological yield (BY), grain yield (GY), harvest index (HI) and 1000-grain weight (GW)	55 accessions / genotypes from the Directorate of Wheat Research, Karnal, India	SSR (S), SAMPL (A) and AFLP (F)	SLA and MRA	One S and two F markers associated with PH; six S, three A and four F for DF; three S and one A for DM; one S, two A and three F for TN; four S, three A and one F for FLA; two S and one F for PL; one A and two F for SL; four S and five F for SP; two S, one A and seven F for FL; two S, one A and six F for GR; one S and three F for BY; three S and one F for GY; four S, one A and one F for HI; two F for GW	Roy et al., 2006
Mulberry (<i>Morus indica</i> L.)	Leaf yield attributing traits including weight of 100 leaves (WL), branches internodal distance (BID), leaf yield (LY), chlorophyll (CH), protein (PR), moisture (MO) and total shoot length (TSL)	44 genotypes collected from different parts of Kerala	ISSR	MRA	Seven markers associated with WL, seven for BID, three for LY, three for CH, five for PR, two for MO and five for TSL	Vijayan et al., 2006
Alfalfa (<i>Medicago sativa</i>)	Forage yield and fall growth	A C0 population generated from the Peruvian \times WISFAL population for several generations	RFLP and SSR	MRA	Five markers associated with forage yield and seven markers for fall growth	Maureira-Butler et al., 2007

development.

Well then, how do we conduct the GRC marker-trait association identification? We should first select optimal accessions from germplasm and evaluate the phenotypic values of interesting traits per accession and then amplify each accession using optimal molecular marker. After then, we do marker-trait association identification by the GRC analysis and then test identification efficiency of markers associated with traits (see below). In this process, marker-trait associations may be screened using the MRA approach: each quantitative trait is treated as a dependent variable and the various marker genotypes (scored as 1 for presence and 0 for absence) as independent

variables. The analysis is based on the model:

$$Y = a + b_1m_1 + b_2m_2 + \dots + b_jm_j + \dots + b_nm_n + d + e$$

which relates the variation in the dependent variable (Y = accession means for a quantitative trait) to a linear function of the set of independent variables m_j , representing the markers. The b_j terms are the partial regression coefficients that specify the empirical relationships between Y and m_j ; d represents the accessions residual which is left after regression and e is the random error of Y that includes environmental variation (Virk et al., 1996; Kar et al., 2008). F-values with P-values between 0.045

Table 1. Continued

Plants	Traits	Materials	Markers	Regression	Markers associated with traits	References
Coconut (<i>Cocos nucifera</i> L.)	Mite resistance (MR)	40 mite-resistant and -susceptible genotypes from different parts of south India	RAPD and SSR	SLA and MRA	SLA: nine SSR and four RAPD markers associated with MR; MRA: six SSR and three RAPD markers associated with MR, a combination of three SSR and two RAPD with 100% association with MR	Shalini et al., 2007
Birch (<i>Betula platyphylla</i>)	Fiber length (FL)	100 individuals	SSR, RAPD and ISSR	MRA	Four SSR markers associated with FL, PB15 of them could identify as high as 75% long fiber individuals; Identification efficiencies of two SCAR markers transferred from RAPD _{BFLI-16} and ISSR _{BFLI-3} markers associated with FL were > 92% for long-fiber length and 76% for short-fiber length respectively	Wang, 2007; Wang et al., 2008; Xia et al., 2008
Oat (<i>Avena sativa</i> L.)	Agronomic traits (grain and groat yield, panicle emergence, plant height and lodging) as well as kernel quality traits (kernel and test weight, screening and groat percent)	One hundred and fourteen varieties of worldwide origin	AFLP	GLM and MLM	23 AFLP markers with many of these affecting multiple traits	Achleitner et al., 2008
Mulberry (<i>Morus</i> ssp.)	Protein content	18 selected accessions from <i>M. indica</i> , <i>M. alba</i> and <i>M. laevigata</i>	ISSR	MRA	Four markers associated with protein content	Kar et al., 2008
Sea buckthorn (<i>Hippophae</i> L.)	Resistance and susceptibility to dried-shrink disease (DSD)	52 accessions from <i>H. rhamnoides</i> and <i>H. salicifolia</i>	ISSR	MRA	Four markers associated with DSD-resistance; four for DSD-susceptibility	Ruan et al., 2009

RAPD: Random amplified polymorphic DNA; MRA: multiple regression analysis; AFLP: amplified fragment length polymorphism; SSR: simple sequence repeat; SAMPL: selective amplification of microsatellite polymorphic loci; SLA: single linear regression analysis; ISSR: inter-simple sequence repeats; RFLP: restriction fragment length polymorphism; GLM: general linear model; MLM: mixed linear model.

and 0.099 were used to enter and remove independent variables from the regression equation, respectively (Affifi and Clark, 1984; Roy and Bargmann, 1957). R^2 denotes the square of r , the correlation coefficient. Selected markers should be further tested with linear curve fitting, using linear models for confirming the significance of β -statistics for each band identified by MRA. Beta can be defined as standardized regression co-efficient = BS_x/S_y , where B is the regression coefficient or slope and S_x and S_y are the standard deviations of independent (x) and dependent (y) variables (Affifi and Clark, 1984; Kar et al., 2008; Ruan et al., 2009). This method provides maximum likelihood estimates of relationships between individual quantitative traits and various markers, and has been successfully

used in some studies shown in Table 1.

FACTORS THAT AFFECT GRC MARKER-TRAIT ASSOCIATION IDENTIFICATION

The success of MAS/QTL based on linkage map is determined by availability of mapping populations and tight linkage between markers and QTLs. Therefore, the essential validation of the GRC marker-trait association identifications should largely depend on the selection of optimal germplasm materials and molecular markers and the test of identification efficiency of markers associated with traits. First, we should know how to select optimal

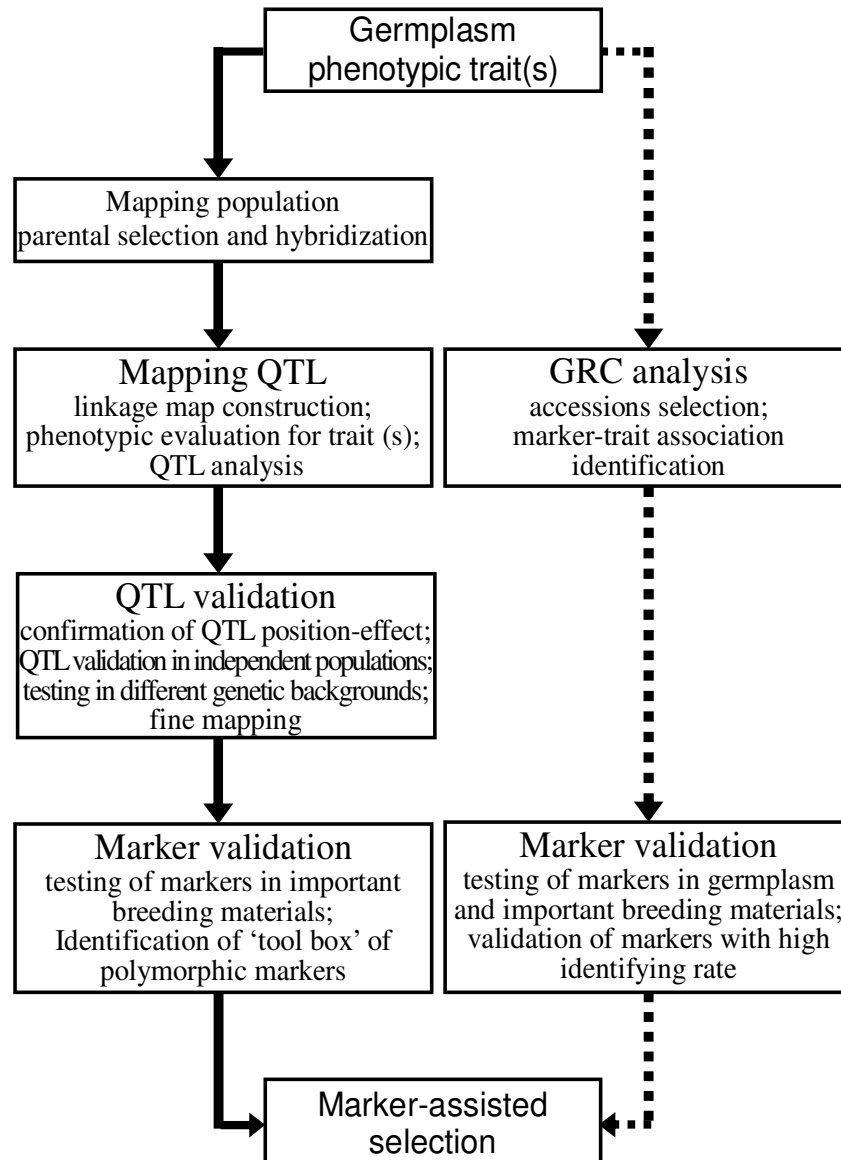


Figure 1. Comparison in marker developments between linkage-based (solid arrowheads, cited from Collard and Mackill, 2008) and germplasm-regression combined (dashed arrowheads) marker-trait association identification.

germplasm as sampling accessions in the GRC association analysis. Based on the schemes of MAS reviewed by Kearsey and Farquhar (1998), Asins (2002), Collard et al. (2005), Mohler and Singrun (2005), Holland (2007), Collard and Machill (2008), Hospital (2009) and this paper, ideal accessions used in GRC marker-trait association identification should have (i) obviously phenotypic difference in interesting traits, (ii) widely genetic diversity and (iii) optimal genetic background. If there are obvious differences in phenotypic traits among different individuals belonging to one genotype or different genotypes of one species, which grow in the same or different regions, these samples are suitable for the GRC analysis, but we

should pay attention to intraspecific genetic diversity. This sampling strategy was used in some GRC association studies (Virk et al., 1996; Roy et al., 2006; Vijayan et al., 2006; Shalini et al., 2007; Wang, 2007; Achleitner et al., 2008). On the other hand, accessions from different species belonging to the same genus in some studies (Kar et al., 2008; Ruan et al., 2009) were used for the GRC analysis. In this case, optimal genetic background should be noticed to avoid the frequency of specific alleles in multiple germplasms that may be too low to detect their effects (Maureira-Bulter et al., 2007).

Secondly, we should know how to select optimal molecular markers in the GRC association analysis.

Traditionally, morphological markers have general demerits that reduce their usefulness. These include delay of marker expression that results in late development of the organism, dominance, deleterious effects, pleiotropy, confounding effects of genes that are unrelated to the gene or trait of interest, but which also affect the morphological marker (epistasis), rare polymorphism and highly influencing environmental conditions. To avoid problems specific to morphological markers, molecular markers with various qualities have been developed (Teixeira da Silva et al., 2005). They are highly polymorphic, with simple inheritance (often co-dominant), and occur abundantly throughout the genome; they are easy and fast to detect, have minimum pleiotropic effect and detection that is not dependent on the developmental stage of the organism. Well then, what markers are optimal for MAS? Important properties of ideal molecular markers associated with traits of interest should include (i) easy recognition of all possible phenotypes (homo and heterozygotes) from all different alleles, such as sequence-related amplified polymorphism (SRAP) and SSR markers, (ii) measurable differences in expression between trait types and/or gene of interest alleles, early in the development of the organism, (iii) no effect on the trait of interest that varies depending on the allele at the marker loci, (iv) low or null interaction among the markers allowing the use of many at the same time in a segregating population and (v) abundance in number and high polymorphism.

However, so far, most types of molecular markers (for example RFLP, RAPD, ISSR, AFLP and SRAP), though nowadays PCR-based, are still too impractical to be used in large-scale MAS schemes due to the complexity of the assay preventing the appropriate automation, insufficient robustness or inadequate level of detected polymorphism (Koeber and Summers, 2003). The most widely used markers in major cereals are SSR markers (Collard and Mackill 2008), which are highly reliable (that is reproducible), have co-dominant inheritance, relatively simple and cheap to use and generally highly polymorphic; but SSR markers require a substantial investment of time and money to develop and plants with developed SSR markers are still very limited, especially for orphan crops and long-juvenile woody species. Due to high polymorphic information content, sequence-tagged site (STS), sequence characterized amplified region (SCAR) that are derived from specific DNA sequences of markers (for example RAPD, RFLP, AFLP, ISSR and SRAP) and single nucleotide polymorphism (SNP) markers, are presently the most appropriate marker class for MAS (Sanchez et al., 2000; Sharp et al., 2001; Mohler and Singrun, 2005; Collard and Mackill, 2008). For orphan crops and woody species that have no genetic information used to develop STS and SNP markers, SCAR marker still will be a potential marker for their MAS in the future.

Finally, the testing of identification efficiency of markers associated traits is one key step for the MAS. Identifi-

fication percentages of two SCAR markers transferred from ISSR_{BFLI-3} and RAPD_{BFLI-16} that are associated with fiber birch length were 76 for short fiber (Xia et al., 2008) and over 92% for long fiber in birch (Wang et al., 2008), respectively. In four SSR markers associated with fiber birch length of birch, PB15_{M3} could identify as high as 75% long fiber individuals (Wang, 2007).

FUTURE PROSPECTS

Although the extent of MAS use will depend on available resources and may be delayed in less-developed countries, the greater adoption of MAS in the future is inevitable (Collard and Mackill, 2008), especially for orphan crops and long-juvenile woody plants. The cost of genotyping (an example of a molecular marker assay) is reducing while the cost of phenotyping is increasing particularly in developed countries, thus increasing the attractiveness of MAS as the development of the technology continues. MAS will then increasingly be applied to obtain improved efficiency and effectiveness in the selection of genotypes with traits that are difficult and expensive to phenotype, for the pyramiding of disease resistance genes in single genotypes and for the carefully directed choice of parental lines in crossing programs allowing a controlled combination of alleles targeted for selection (Mohler and Singrun, 2005).

The tight linkage between markers with genes/QTLs is critical to the success of MAS. Identification of genes/QTLs at distances over 2 cM from the closest markers is hardly suitable either for MAS or for identification/cloning of functional genes (Mohler and Singrun, 2005), however, in some cases, recombination occurs between the marker and genes/QTL due to loose linkage (Collard and Mackill, 2008). In addition, even as Virk et al. (1996) suggested that since the success of MAS programs depends exclusively on the extent of genetic linkage between markers and the relevant loci such as QTL, there is no reason why the same principles cannot be applied to either natural populations or genetic resources generally, assuming that similar associations are observed between marker loci and the various allelomorphic forms of QTLs and that the basis of these is in fact genetic linkage. Therefore, the GRC molecular marker-trait association identification will provide several potentials for the improvement of quantitative traits in orphan crops and long-juvenile woody plants, including marker-assisted evaluation of breeding material or germplasm, early selection of desirable traits and effective selection of putative parents for producing populations to map QTLs for a particular trait or cloning particularly functional genes.

First of all, use of new crops and woody plants, which can grow in the areas of mounting water scarcity, environmental degradation, increasing pollution and inevitable emergence of new biotypes of pathogens and pests, is increasing, accompanying the requirement of increased

crop production from a rising global population but the declining rate of increase in crop yields (Collard and Mackill, 2008). However, the linkage-based identification of markers associated with genes/QTLs often needs to be too long-time for these orphan crops and woody plants with high heterozygosity. For example, in general at least 9 - 10 years are required for mulberry to express all its traits to its full potential (Vijayan et al., 2006). Meantime, selection of individual genotypes for mapping of traits in bi-parental populations will have sampled only a small portion of the resident variation within the original germ-plasm and may have reduced the odds of finding regions affecting the traits of interest (Maureira-Bulter et al., 2007). Hence, for orphan crops and long-juvenile woody plants with strong tolerance to harsh environments, in which most of them have no genetic information such as linkage maps and Quantitative Trait Loci, marker-assisted evaluations of breeding material or germplasm by the GRC association studies should allow MAS to become more widely applicable for their breeding programs, especially for screening germplasm with desirable traits.

Secondly, the early selection of desirable traits that is expressed late in plant development, like fruit and flower features or adult characters with a juvenile period, so that it is not necessary to wait for the organism to become fully developed before arrangements can be made for propagation. For long juvenile and long-lived woody plants with heterozygosity, progeny selection is always a difficult process as the expression of most of the agronomical traits varies according to the stages of development (Vijayan et al., 2006). Hence, the GRC molecular marker-traits association identification could be of great use for the breeder to identify promising seedling/hybrids at an early stage of these plants. For example, sea buckthorn is highly heterozygous and has a juvenile period from 3 - 5 years. Dried-shrink disease (DSD) is a dangerous pathogen that destroys this species and halts commercial production, but it is always a difficult process for sea buckthorn to select materials with DSD-resistance in conventional breeding programs. This is because DSD infects plants that are a minimum of three years old, whereas all plants infected by DSD are more than 3 years old and in general at least 3 - 5 years are required for the sea buckthorn plant to express symptom of DSD to its full potential (Ruan et al., 2008). The QTL linkage map is unknown for sea buckthorn, thus construction of mapping populations (even F_1 crosses) is time-consuming and labor-intensive. Hence, the ISSR markers associated with DSD resistance identified by GRC analysis (Ruan et al., 2009) could be used to eliminate futureless seedling/hybrids at an early stage: if the markers associated with high disease index were identified in the seedlings, it should fall into disuse to prevent great environmental and commercial loss because of the death of sea buckthorn plantations from DSD-infection.

Finally, if genetic linkage is the main cause of the associations, then one benefit of obtaining information about

molecular markers and quantitative traits could be for use in more efficiently selecting putative parents for producing populations to map QTL for a particular trait and for cloning particularly functional genes. For example, mapping population used to map QTL of DSD-resistance in sea buckthorn may be constructed by selection parents which genomes cover the molecular markers associated with DSD-resistance and susceptibility respectively.

In conclusion, the apparent advantages of the GRC marker-trait association identification are (i) that this could allow the detection of QTL that varies across a wide spectrum of biodiversity rather than just between two planned parental genotypes; (ii) that QTL for any quantitative trait can be studied in the same investigation and (iii) that this requires less inputs of time, labor and financial resources, compared to the linkage-based QTL identification. The GRC marker-trait association identification will play an important role in plant MAS/QTL breeding programs, especially in orphan crops and long-juvenile woody plants with heterozygosity when no other genetic information such as linkage maps and Quantitative Trait Loci are available. However, in this review we also do not try to minimize successes of molecular genetics by laying emphasis on quantitative trait loci (QTL) identification based on the linkage-map, in which many major genes have been identified and successfully used in the MAS for the improvements of quantitative traits in major crops (Hansen et al., 2008; Hospital, 2009; Jannink et al., 2009).

ACKNOWLEDGEMENT

This work was funded by the Keystone Project of the Agriculture Science and Technology Research Item of the Liaoning Scientific and Technological Committee (No. 2007207005).

REFERENCES

- Achleitner A, Tinker NA, Zechner E, Buerstmayr H (2008). Genetic diversity among oat varieties of worldwide origin and associations of AFLP markers with quantitative traits. *Theor. Appl. Genet.* 117: 1041-1053.
- Affifi AA, Clark V (1984). Computer-aided multivariate analysis. Van Nostrand Reinhold, New York.
- Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005). Cytokinin oxidase regulates rice grain production. *Science*, 309: 741-745.
- Asins MJ (2002). Present and future of quantitative trait locus analysis in plant breeding. *Plant Breed.* 121: 281-291.
- Barloy D, Lemoine J, Abelard P, Tanguy AM, Rivoal R, Jahier J (2007). Marker-assisted pyramiding of two cereal cyst nematode resistance genes from *Aegilops variabilis* in wheat. *Mol. Breed.* 20: 31-40.
- Benchimol LL, de Souza CL, de Souza AP (2005). Microsatellite-assisted backcross selection in maize. *Genet. Mol. Biol.* 28: 789-797.
- Bouchez A, Hospital F, Causse M, Gallais A, Charcosset A (2002). Marker-assisted introgression of favorable alleles at quantitative trait loci between maize elite lines. *Genetics*, 162: 1945-1959.
- Chatterjee SN, Mohandas TP (2003). Identification of ISSR markers associated with productivity traits in silkworm, *Bombyx mori* L. *Genome*, 46: 438-447.

- Chatterjee SN, Pradeep AR (2003). Molecular markers (RAPD) associated with growth, yield and origin of the silkworm, *Bombyx mori* L. in India. *Russ. J. Genet.* 39: 1365-1377.
- Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142: 169-196.
- Collard BCY, Mackill DJ (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Phil. Trans. R. Soc. B.* 363: 557-572.
- Davies J, Berzonsky WA, Leach GD (2006). A comparison of marker-assisted and phenotypic selection for high grain protein content in spring wheat. *Euphytica* 152:117-134.
- Ender M, Terpstra K, Kelly JD (2008). Marker-assisted selection for white mold resistance in common bean. *Mol. Breed.* 21: 149-157.
- Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schuler K (2004). Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Mol. Breed.* 13: 93-102.
- Hansen BG, Halkier BA, Kliebenstein DJ (2008). Identifying the molecular basis of QTLs: eQTLs add a new dimension. *Trends Plant Sci.* 13: 72-77.
- Helguera M, Khan IA, Kolmer J, Lijavetzky D, Zhong-qi L, Dubcovsky J (2003). PCR assays for the Lr37-Yr17-Sr38 cluster of rust resistance genes and their use to develop isogenic hard red spring wheat lines. *Crop Sci.* 43: 1839-1847.
- Holland JB (2007). Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* 10: 156-161.
- Hospital F (2009). Challenges for effective marker-assisted selection in plants. *Genetica*, 136: 303-310.
- Jannink JL, Moreau L, Charmet G, Charcosset A (2009). Overview of QTL detection in plants and tests for synergistic epistatic interactions. *Genetica*, 135: 255-236.
- Kar PK, Srivastava PP, Awasthi AK, Urs SR (2008). Genetic variability and association of ISSR markers with some biochemical traits in mulberry (*Morus* spp.) genetic resources available in India. *Tree Genet. Genomes* 4: 75-83.
- Kearsey MJ, Farquhar AGL (1998). QTL analysis in plants: where are we now? *Heredity*, 80: 137-142.
- Knoll J, Ejeta G (2008). Marker-assisted selection for early-season cold tolerance in sorghum: QTL validation across populations and environments. *Theor. Appl. Genet.* 116: 541-553.
- Koebner R, Summers R (2003). 21st century wheat breeding: selection in plots or detection in plates? *Trends Biot.* 21: 59-63.
- Liao CY, Wu P, Hu B, Yi KK (2001). Effects of genetic background and environment on QTLs and epistasis for rice (*Oryza sativa* L.) panicle number. *Theor. Appl. Genet.* 103: 104-111.
- Maureira-Butler IJ, Udall JA, Osborn TC (2007). Analyses of a multi-parent population derived from two diverse alfalfa germplasm: testcross evaluations and phenotype-DNA associations. *Theor. Appl. Genet.* 115: 859-867.
- Mishra RK, Sen-Mandi S (2004). Molecular profiling and development of DNA marker associated with drought tolerance in tea clones growing in Darjeeling. *Curr. Sci.* 87: 60-66.
- Misztal I (2006). Challenges of application of marker assisted selection-a review. *Animal Sci. Papers Rep.* 24: 5-10.
- Mohler V, Singrun C (2005). General Considerations: Marker-Assisted Selection. In: Lorz H, Wenzel G (eds.), *Molecular Marker Systems in Plant Breeding and Crop Improvement* (Section III), Springer, Berlin Heidelberg, pp. 305-317.
- Obert DE, Skinner DZ, Stuteville DL (2000). Association of AFLP markers with downy mildew resistance in autotetraploid alfalfa. *Mol. Breed.* 6: 287-294.
- Pradeep AR, Jingade AH, Urs RS (2007). Molecular markers for biomass traits: association, interaction and genetic divergence in silkworm *Bombyx mori*. *Biomarker Insights* 2: 197-217.
- Ren ZH, Gao JP, Li LG, Cai XL, Huang W, Chao DY, Zhu MZ, Wang ZY, Luan S, Lin HX (2005). A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* 37: 1141-1146.
- Ribaut JM, Hoisington DA (1998). Marker assisted selection: new tools and strategies. *Trends Plant Sci.* 3: 236-239.
- Roy JK, Bandopadhyay R, Rustgi S, Balyan1HS, Gupta PK (2006). Association analysis of agronomically important traits using SSR, SAMPL and AFLP markers in bread wheat. *Curr. Sci.* 90: 683-689.
- Roy SN, Bargmann RE (1957). Tests of multiple independence and the associated confidence-bounds. North Carolina Institute of Statistics Mimeograph Series No. 175.
- Ruan CJ, Li H, Mopper S (2009). Characterization and identification of ISSR markers associated with resistance to dried-shrink disease in sea buckthorn. *Mol. Breed.* 24: 255-268.
- Ruan CJ, Zhang J, Zhou ZZ (2008). Symptom & pathogenous law of dried-shrink disease of *Hippophae* L. and primary selection of resistant varieties. *J. Dalian National Univ.* 10: 390-393 (in Chinese with an English abstract).
- Sanchez AC, Brar DS, Huang N, Li Z, Khush GS (2000). Sequence tagged site marker-assisted selection for three bacterial blight resistance genes in rice. *Crop Sci.* 40: 792-797.
- Shalini KV, Manjunatha S, Lebrun P, Berger A, Baudouin L, Pirany N, Ranganath RM, Prasad DT (2007). Identification of molecular markers associated with mite resistance in coconut (*Cocos nucifera* L.). *Genome*, 50: 35-42.
- Sharp PJ, Johnston S, Brown G, McIntosh R, Pallotta MA, Carter M, Bariana HS, Khatkar S, Lagudah ES, Singh RP, Khairallah M, Potter R, Jones MGK (2001). Validation of molecular markers for wheat breeding. *Aust. J. Agric. Res.* 52: 1357-1366.
- Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39: 623-630.
- Srivastava PP, Kar PK, Awasthi AK, Urs SR (2007). Identification and association of ISSR markers for thermal stress in polyvoltine silkworm *Bombyx mori*. *Rus. J. Genet.* 43: 858-864.
- Steele KA, Price AH, Shashidhar HE, Witcombe JR (2006). Marker-assisted selection to introgress rice QTLs controlling root traits into an Indian upland rice variety. *Theor. Appl. Genet.* 112: 208-221.
- Teixeira da Silva JA, Nhut DT, Giang DTT, Rashid SZ (2005). The phylogeny, breeding and ecology of trees using molecular markers. In: Thangadurai D, Pullaiah T, Tripathy L (Eds.), *Genetic Resources and Biotechnology* (Vol. III), Regency Publications, New Delhi, pp. 279-308.
- Thomas W (2003). Prospects for molecular breeding of barley. *Ann. Appl. Biol.* 142: 1-12.
- Vijayan K, Srivatsava PP, Nair CV, Awasthi AK, Tikader A, Sreenivasa B, Urs SR (2006). Molecular characterization and identification of markers associated with yield traits in mulberry using ISSR markers. *Plant Breed.* 125: 298-301.
- Virk PS, Ford-Lloyd BV, Jackson MT, Pooni HS, Clemeno TP, Newbury HJ (1996). Predicting quantitative variation within rice germplasm using molecular markers. *Heredity*, 76: 296-304.
- Wang D, Wei ZG, Yang CP, Liu GJ (2008). Analysis and identification of SCAR molecular markers associated with birch fiber length trait. *J. Forest Res.* 19: 288-292.
- Wang YQ (2007). Study on *Betula platyphylla* long fiber trait by SSR marker technology. Dissertation for the Degree of Master, Northeast Forestry University (in Chinese with an English abstract).
- Wilde F, Korzun V, Ebmeyer E, Geiger HH, Miedaner T (2007). Comparison of phenotypic and marker-based selection for *Fusarium* head blight resistance and DON content in spring wheat. *Mol. Breed.* 19: 357-370.
- Wright AJ, Mowers RP (1994). Multiple regression for molecular-marker, quantitative trait data from large F₂ populations. *Theor. Appl. Genet.* 89: 305-312.
- Xia DA, Wei ZG, Yang CP, Liu GJ (2008). Analysis of ISSR and SCAR Markers associated with birch fiber length trait. *J. Northeast For. Univ.* 39: 1-4 (in Chinese with an English abstract).
- Yonash N, Heller ED, Hillel J, Cahaner A (2000). Detection of RFLP markers associated with antibody response in meat-type chickens: haplotype/genotype, single-band and multiband analyses of RFLP in the major histocompatibility complex. *J. Hered.* 91: 24-30.
- Zhang J, Li X, Jiang G, Xu Y, He Y (2006). Pyramiding of Xa7 and Xa21 for the improvement of disease resistance to bacterial blight in hybrid rice. *Plant Breed.* 125: 600-605.