Designing crossing and selection strategies to combine diagnostic markers and quantitative traits

Chapman SC[1], Wang J[2], Rebetzke GJ[3], Bonnett DG[3]
[1]CSIRO Plant Industry, St. Lucia, QLD 4067, Australia. [2] CIMMYT, c/o Institute of Crop Sciences, CAAS, Beijing 100081, China,[3]CSIRO Plant Industry, Black Mountain, ACT 2601, Australia

ABSTRACT

Molecular markers are now in common use, particularly where diagnostic markers (i.e. allele-based functional markers or closely linked markers) exist. Diagnostic markers are now available for more than 20 traits in wheat. While many markers for quantitative trait loci (QTL) have been published, few are actually deployed in breeding. The challenge for wheat breeders is how to implement marker systems with phenotypic selection in the creation of new parental lines and progeny (target genotypes). For simple procedures and small numbers (<3) of unlinked genes, population sizes and assays can be calculated from population genetic theory. In previous work ((Wang *et al.* 2007), we showed that simulation was required to find an efficient crossing and selection scheme (i.e. low-cost, small population size) to combine nine diagnostic markers donated by three parents. The genes coded for height (Rht-B1, Rht-D1, Rht-8), grain quality (Glu-B1, Glu-A3), tillering (tin) and disease (Sr2, Cre1, VPM).

The Rht-8 gene increases coleoptile length. Based on published work, we assigned additive effect values for six additional QTL for coleoptile length and used this as a quantitative trait in simulation. From 1000 simulations of a cross between a conventional semi-dwarf wheat (Sunstate) and a long coleoptile donor (HM14BS) carrying the Rht8 allele, we targeted selection of DH lines which had the correct alleles for six major genes and with long coleoptiles (>120 mm). After elimination of less efficient backcross strategies, we created the 1000 DH lines from the F_1. Selection in the DH returned only 2.2 target genotypes from the 1000 lines screened. We then created an F_2 of either 100 (option 2) or 200 (option 3) lines. These were screened for the six major genes (F_2 enhancement selection), and then we created either 900 or 800 DH lines for the two options, and screened these for the genes and long coleoptiles. Options 2 and 3 resulted in an average of 13 or 19 DH lines from the 1000 lines screened, with up to 100 target genotypes found in some simulations. Further simulations are being undertaken to determine more efficient ways of multi-plexing markers and integrating with the phenotypic screening of coleoptile length.

INTRODUCTION

Examples of marker-based applications include; introgression of traits from donor ('unadapted' lines) into parental germplasm; broadening the genetic base of a crop (Xu *et al.* 2004); selection of parental combinations, based on marker profiles (Wang *et al.* 2005); selection of cross progeny during early and late generations of selfing and evaluation (Eagles *et al.* 2001) and fast-throughput recurrent selection for QTL, termed 'mapping-as-you-go' (Podlich *et al.* 2004).

Markers for single gene/single trait applications have been used in wheat breeding programmes in Australia for over 10 years, e.g. (Ogbonnaya *et al.* 2001) and (Eagles *et al.* 2001)). In general, their use has been in introgressing into breeding lines (in BCF_1) and in screening progeny in early (F_2) and later generations of evaluation. With the advent of higher density maps, and increasing effort in phenotyping, there is a greater need to deploy together both 'functional' markers (i.e. associated with the sequence of a specific allele of a gene) with markers that are have been associated with QTL, and for which specific alleles are not yet identified. Strategies for efficient pyramiding of alleles at multiple marker loci have been reported ((Bonnett *et al.* 2005); (Kuchel *et al.* 2007); (Wang *et al.* 2007)). Extension to selection for markers of smaller genetic effect will increase numbers of loci under selection to increase population sizes particularly if marker selection must be balanced with phenotypic selection.

Long coleoptiles allow deeper planting and contribute to improved establishment and yield in rainfed wheat-growing areas. Characteristics of coleoptile length include high heritability and major QTL have been located for the trait across environments ((Rebetzke *et al.* 2001), (Rebetzke *et al.* 2007)). In order to understand and commence development of strategies for selection of polygenic traits, we examined coleoptile length in wheat as a quantitatively-inherited trait under polygenic control. In this paper, we investigate a realistic scenario in breeding where MAS for major genes and genes of minor effect ('QTL') is likely to be integrated with conventional phenotypic selection.

To begin to develop suitable methods of QTL selection for polygenic traits, we examine coleoptile length as an important polygenic, quantitatively inherited trait. In this paper, we demonstrate how to integrate MAS including major genes and QTL into conventional phenotypic selection using the breeding procedure for drought tolerance in wheat as an example. Our breeding objectives are to combine six major genes for disease resistance and high grain quality, and at the same time produce long coleoptile lines.

MATERIALS AND METHODS

HM14BS (P1) is an F6-derived line from a cross between the tall (rht8), long coleoptile cv. Halberd and the semidwarf (Rht8), short coleoptile cv. Mara. This

dwarfing gene is gibberelin-sensitive, and so will produce long coleoptiles when planted deep. Sunstate (P2) is an Australian wheat cultivar, carrying the Green Revolution, reduced height gene Rht-D1b. The coleoptiles are gibberelin-insensitive (not sensitive to darkness) and so will prematurely initiate leaf expansion if planted too deep, resulting in death. Coleoptile length is assessed by sowing seed 35 to 40mg in size at a depth of 20mm in well-watered, deep trays (600×300×120mm) containing a fertile, compost-based potting mix. Lines are replicated and grown in the dark in growth cabinets at constant temperatures of 15 and/or 19°C. for 200°Cd.

The six marker-linked genes (Table 1) are currently being used in selection in some Australian breeding programs to develop new lines for commercial release and are summarised by (Wang *et al.* 2007) Genotypes of the two parents at six target loci including two reduced height loci are known (Table 1).

Table 1 Six major genes and their genotypes for two parents.

Gene symbol	*Rht-D1*	*Rht8*	*Sr2*	*VPM*	*Glu-B1*	*Glu-A3*
Chrom.os.	4DS	2DL	3BS	7DL	1BL	1AS
Marker type	Codom	Codom	Codom	Dom	Codom	Codom
Distance marker and gene (cM)	0.0	0.6	1.1	0.0	0.0	0.0
HM14BS	*Rht-D1a*	*Rht8*	*sr2*	*vpm*	*Glu-B1a*	*Glu-A3c*
Sunstate	*Rht-D1b*	*rht8*	*Sr2*	*VPM*	*Glu-B1i*	*Glu-A3b*
Target genotype†	*Rht-D1a*	*Rht8*	*Sr2*	*VPM*	*Glu-B1i*	*Glu-A3b*

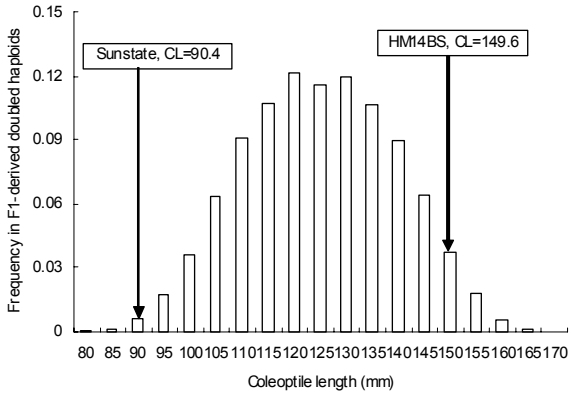

Figure 1 Distribution of coleoptile length in F1-derived DH population based on the genetic model

The reduced height gene Rht-D1 has a large pleiotropic effect on coleoptile length. In this practical example, we attempt to transfer knowledge about QTL from a mapping to a breeding population. Therefore, the QTL effects in (Rebetzke *et al.* 2007) were distributed between the parents by assuming that HM14bS inherited many of the long coleoptile alleles from Halberd and Sunstate contained similar coleoptile alleles as Cranbrook. In that study, the reduced height gene Rht-D1 explained 54% of total additive genetic variance in CL and six other QTL accounted for c. 33% of the additive genetic variance (Table 2). We assumed that there were eight unmapped QTL explaining the remaining 13% of the variance, and distributed these randomly between the parents, resulting in the genetic model shown in Figure 1. The two parents have similar plant heights but HM14BS has a longer CL than Sunstate (cf. 140 mm vs. 100 mm for evaluation at 15°C). In selection, for this cross, lines were considered to have a 'long coleoptile' trait, if CL exceeded 120mm. For simulation of plant height and coleoptile length, heritability in the narrow-sense was set at 0.76 (Rebetzke *et al.* 2007).

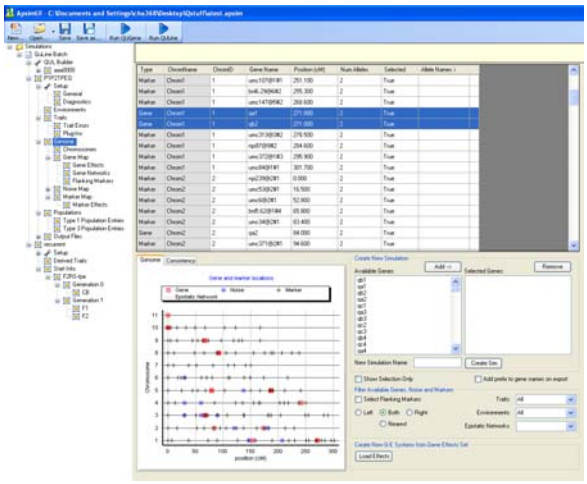

Figure 2 Screenshot of genome editing interface

Qu-Line is a genetics and breeding simulation tool that can integrate various genes with multiple alleles operating within epistatic networks and differentially interacting with the environment interaction, and to predict the outcomes from a specific cross following the application of a real selection scheme The software is available from The University of Queensland, Australia http://www.uq.edu.au/lcafs/qugene/. It allows the interactive entry of all marker and gene information from data on maps, genotypes and population composition (Figure 2).

Under a fixed resource, i.e. a total of 1000 individuals in early segregating generations such as F2 and F3 can be grown, genotyped and phenotyped, and 200 advanced lines (DH lines derived from F1, F2 or F3, depending on the selection scheme, in this study). Our breeding objective is to select 20 DH lines combining the six major resistance and quality genes, and having coleoptile length more than 120mm. Simulations were repeated 1000 times.

RESULTS AND DISCUSSION

In a preliminary analysis, we compared DH lines from F_1 and backcross strategies. For the six major lines, a backcross to HM14BS ($P1BC_1$) resulted in the least number of target DH lines (2.2 lines per 1000),

compared to 19.6 and 15.5 in 1000 DH lines derived from $P2BC_1$ and F_1, respectively (data not shown). When coleoptile length was also considered DH lines from the F1 had the population highest target frequency and so backcross was not considered further.

From the F1 derived lines, 15.5 per 1000 have the target genotype for the six major genes (including the major gene for height and coleoptile length), and 74.1 lines have coleoptile length greater than HM14BS (Option 1, Table 2). However, when both targets were considered, only 2.4 DH lines would be found.

Table 2 Distribution characteristics (min, max, mean from 1000 simulations) of target genotypes selections. MG = number of lines having all six major genes; CL = number of lines > 120mm; MC = both criteria met

	Option 1 Selection was conducted with 1000 DH lines			Option 2 Selection was conducted with 900 DH lines derived from F2 individuals after enhancement selection			Option3 Selection was conducted with 800 DH lines derived from F2 individuals after enhancement selection		
	MG	CL	MC	MG	CL	MC	MG	CL	MC
Min	5	52	0	33	19	0	41	0	1
Max	31	99	7	178	262	66	316	364	100
Mean	16	74	2.4	79	89	13.0	117	133	19.3

Given that we have markers for six major genes, we can conduct selection in F2, and then use doubled haploids to derive inbred lines. In this way, the frequency of targets can be increased. We considered two additional options: enhancement selection (for the markers of all six genes) applied on 100 F2 individuals, and then 900 DH lines derived from selected F2 individuals, and enhancement selection applied on 200 F2 individuals, and then 800 DH lines derived from selected F2 individuals. Considering the extraction of DNA samples is the major cost for applying MAS, we assume that the two options will result in similar expenses, if not same. In option 2, on average, 78.6 DH lines can be selected having the target genotype of the six major genes, 88.9 DH lines having CL > 120 mm, and 13.0 DH lines meet both breeding objectives (Table 2). For option 3, on average, 116.7 DH lines can be selected having the target genotype of the six major genes, 133.0 DH lines having CL > 120 mm, and 19.3 DH lines meet both breeding objectives. Based on 1000 simulations, the worst case for option 3 is that only 1 DH line is found, while the best case is that 100 of the 1000 DH lines meets the criteria.

ACKNOWLEDGEMENTS

Parts of this research has been funded or in collaboration with the CGIAR Generation Challenge Program and the Australian Grains Research and Development Corporation.

REFERENCES

Bonnett DG, Rebetzke GJ, Spielmeyer W (2005) Strategies for Efficient Implementation of Molecular Markers in Wheat Breeding. *Molecular Breeding* **15**, 75-85.

Eagles HA, Bariana HS, Ogbonnaya FC, Rebetzke GJ, Hollamby GJ, Henry RJ, Henschke PH, Carter M (2001) Implementation of markers in Australian wheat breeding. *Australian Journal of Agricultural Research* **52**, 1349-1356.

Kuchel H, Fox R, Reinheimer J, Mosionek L, Willey N, Bariana H, Jefferies S (2007) The successful application of a marker-assisted wheat breeding strategy. *Molecular Breeding* **20**, 295-308.

Ogbonnaya FC, Subrahmanyam NC*, et al.* (2001) Diagnostic DNA markers for cereal cyst nematode resistance in bread wheat. *Australian Journal of Agricultural Research* **52**, 1367-1374.

Podlich DW, Winkler CR, Cooper M (2004) Mapping As You Go: An Effective Approach for Marker-Assisted Selection of Complex Traits. *Crop Sci* **44**, 1560-1571.

Rebetzke GJ, Appels R, Morrison AD, Richards RA, McDonald G, Ellis MH, Spielmeyer W, Bonnett DG (2001) Quantitative trait loci on chromosome 4B for coleoptile length and early vigour in wheat (Triticum aestivum L.). *Australian Journal of Agricultural Research* **52**, 1221-1234.

Rebetzke GJ, Ellis MH, Bonnett DG, Richards RA (2007) Molecular mapping of genes for Coleoptile growth in bread wheat (Triticum aestivum L.). *Theoretical and Applied Genetics* **114**, 1173-1183.

Wang J, Chapman SC, Bonnett DG, Rebetzke GJ, Crouch J (2007) Application of Population Genetic Theory and Simulation Models to Efficiently Pyramid Multiple Genes via Marker-Assisted Selection. *Crop Sci* **47**, 582-588.

Wang J, Eagles HA, Trethowan R, Van Ginkel M (2005) Using Computer Simulation of the Selection Process and Known Gene Information to Assist in Parental Selection in Wheat Quality Breeding. *Australian Journal of Agricultural Research* **56**, 465-473.

Xu Y, Beachell H, McCouch SR (2004) A Marker-Based Approach to Broadening the Genetic Base of Rice in the USA. *Crop Sci* **44**, 1947-1959.