



*Proceedings of the 10th Australian Barley Technical Symposium,
2001*

Marker assisted selection - where to now?

Andrew Barr, Jason Eglinton, Peter Langridge, Pat Warner and Ken Chalmers,

Department of Plant Science, Waite Campus, Adelaide University, Glen Osmond,
South Australia.

Introduction

The Australian barley industry has invested in molecular marker technology for nearly a decade through work conducted by ARC, host organizations and since 1997, via the "National Barley Molecular Marker Program"(NBMMMP), funded by the GRDC. As the NBMMMP enters its final year, it is timely to examine where marker technology has succeeded and failed and where profitable new directions might lead.

Marker development and implementation - pre 2001

The philosophy developed within the NBMMMP in marker development and implementation can be divided into the following steps:

- Identify parents differing in the traits of interest
- Develop a population of plants segregating for the traits of interest (using doubled haploids or single seed descent)
- Screen the population for the traits of interest
- Construct linkage maps of the cross, use bulked segregant analysis, or genomic fingerprinting
- Identify molecular markers that co-segregate with the traits of interest
- Test the applicability and reliability of the markers in predicting the traits in related families (also referred to as marker validation)
- Produce clear and simple protocols for assaying the markers
- Modify breeding strategy to optimise use of Marker Assisted Selection (MAS) relative to alternative selection techniques
- Implement into the breeding programs

There are now 10 major mapping populations and 44 minor populations (for bulked segregant analysis) within the NBMMP as well as six populations developed outside the NBMMP (five at the Waite and one at Wagga), so there is a very rich genetic and phenotypic resource available to the Australian industry. As a result of these combined investigations, it is now possible to undertake serious marker assisted selection tasks within breeding programs. Several examples of these strategies follow.

Single gene transfer

Following successful marker development, marker assisted selection has been used to transfer single genes including *Ha2* (resistance to cereal cyst nematode), spot form net blotch (*Rpt 4*) and *Yd2* (resistance to BYDV) into advanced breeders lines now entered in Stage 4 trials. Other breeding programs have used MAS to incorporate genes for mildew resistance (*mlo*), boron tolerance, manganese efficiency (*Mel1*), aluminium tolerance, leaf rust, stem rust, semi-dwarfism (*sdw*) and cereal cyst nematode (*Ha4*). Marker assisted selection is especially attractive compared to 'conventional assays' for recessive alleles such as *mlo*, where heterozygous carriers of the trait cannot be detected using phenotypic selection.

This technology is now well established and can offer even more potential benefits by

- Simultaneous selection for many traits from the donor parent. This is usually not practical with other selection systems.
- Selection for the recurrent parent genotype over the whole genome but especially around the gene (5) of interest, where a small introgression segment is often highly desirable.

QTL tracking

MAS is an especially valuable tool in many fast-track strategies. For instance, in a single seed descent program initiated in 1998 to rapidly introduce resistance to spot form net blotch (SFNB) and cereal cyst nematode from Keel into Gairdner, we used MAS and bioassays in large populations of the BC₁F₁ and BC₁F₂ to select for *Rpt4* - the gene conferring seedling resistance to spot form net blotch. BC₁F₄ individuals were multiplied over summer and placed into Stage 1 field trials in 2000. Following phenotypic selection for SFNB and CCN resistance and 'Gairdner' plant type, the remaining 43 individuals were tested for *Yd2*, *Rpt4*, *Bmy1* (Gairdner carries the SD1 allele for β-amylase) and three malt extract QTL (1H, 2HL, 5H). Four individuals were promoted to Stage 4 trials in 2001, based on the marker profile and their field performance in 2000. In this cross, the Gairdner allele for malt extract on 1H was the most significant, providing 2% higher extract than the Keel allele. This strategy would have been greatly assisted by markers for CCN resistance and the 'second' gene in SFNB resistance. Further, it would have been useful to have access to a wider range of SSR markers and the ability to deploy them cost effectively. Whilst, this project started with a BC₁F₁ population of over 150 individuals per generation, there are still too few selected lines by the BC₁F₇. Whether the original aims will be achieved

remains to be seen - further backcrosses to Gairdner may be required to fully recapture the Gairdner phenotype for yield, adaptation and quality.

The next phase in the Gairdner 'defect elimination' is to introduce genes for boron tolerance, leaf rust resistance and the SD3 allele for thermostable β -amylase. The crossing strategy involves merging four streams, as follows

CCN + SFNB stream	Boron stream	SD3 stream	Leaf rust stream
Gairdner/Keel//Gairdner	Gairdner/DH115//Gairdner	Gairdner/SD3//Gairdner	Complex cross of Gairdner and Fanfare

By the time the seed of this cross is available, it is hoped that markers for the 'missing' genes will be available and that it will be practically possible to select for the Gairdner background genotype.

Analysis of key malt quality traits

Malt extract is a complex trait conditioned by up to 24 genes (Hayes and Jones, 2000) and subject to considerable G x E. QTL analysis of three Waite and several NBMMP populations have greatly assisted our understanding of this trait. The most important loci detected in Australian studies are closely linked to *Ebmac0501*(1H), *Bmag0378 - Bmag0125* (2HL), *HVM 36*(2HS) and an extended region of 5H, where there may be one to three QTL. Further, the underlying basis of G x E for extract is now clearer. For instance, differences in the Alexis x Sloop population show that regions on chromosome 1H, 2HL and 5H from Alexis are most commonly associated with high extract. However, from two sites in WA, the Sloop allele at the *sdw* locus on 3H conferred high malt extract, whereas this was not associated with high extract lines at other sites. The effect of the 1H, 2HL and 5H loci also differ across sites.

NIR is now routinely used in most of the Australian breeding programs to select for malt extract (on whole grain) in early generation material. The heritability of NIR predications is high (Barr *et al.*, 1999) but it had not been established which QTL are selected. When applied to mapping populations, it is now clear that NIR does not 'detect' all of the QTL for malt extract, such as that on 1H and only intermittently on 5H, which can be seen when the same samples are micromalted and analysed in the traditional way with EBC/IOB methods. The QTL on 2HS and 2HL are evident from both NIR predictions from whole grain and micromalt analysis.

New directions

New methods of establishing marker trait associations

The principle techniques used to identify marker trait associations in crop species have been based around the construction of linkage maps or the use of Bulk Segregant Analysis. Both techniques involve the production and phenotyping of special populations. The populations are generally developed from two varieties that show a major difference in the traits targeted for mapping. This leads to several problems:

- There is usually a high cost associated with phenotyping, particularly for traits requiring extensive field trials or complex analysis, such as many aspects of processing quality and yield. Consequently, the size of the mapping population and the number of replicates and sites in phenotyping experiments is often limited, reducing the sensitivity of some of the analyses.
- The lines (varieties) used to construct the populations are often out-of-date by the time the marker/trait information is available. Many marker development projects for annual crops are using populations that were established five or more years before the marker development work. This reduces the value of the information gathered and the scope of its implementation.
- The structure of the populations limits the types of traits that can be mapped and many of the subtleties of adaptation cannot be analysed.
- Mapping is frequently restricted to known traits for which a well-defined bioassay is available.

These limitations in existing mapping strategies can be addressed through association or linkage disequilibrium (LD) mapping. LD mapping is based on seeking associations between phenotype and allele frequencies at the population level. Equilibrium is seen in large populations over many generations where there is no selective advantage, or disadvantage, associated with a particular allelic combination in a region of the genome. Disequilibrium appears where selective pressure increases, or decreases, the frequency of particular alleles or allelic combinations. It can be measured through an estimation of changes in allele frequencies as a result of the selective pressure. In this way particular regions of the genome can be associated with particular traits.

There are four potential advantages of this approach in mapping genes in crop species:

- It provides a new perspective for trait mapping, because it uses population structures (based largely around pedigree) and phenotypic data that differ from those used for full map construction or BSA. Consequently, we can expect to see new marker trait associations and targets for more detailed analysis.

- LD mapping also provides detailed fingerprinting information on a large number of lines and varieties and this information will be valuable in several of the breeding strategies outlined below.
- The LD method uses real breeding populations, the material is diverse and relevant and the most important genes (for adaptation etc) should be segregating in such populations. The breeder is also integrally involved in the process and this may lead to improved rate and efficiency of validation and adoption. Plant breeders are often reluctant to grow and assess a huge number of lines with little or no potential for direct commercial outcome such as required for complete map construction. The advantage of LD mapping to the breeder is that mapping and commercial variety development is conducted simultaneously.
- Pattern analysis of marker data might detect complex combinations (even epistatic interactions) between alleles at several loci, which underlie the superior individuals in a breeding population. This might prove difficult to isolate and validate using the full mapping approach.

On the other hand, LD mapping has several potential disadvantages. It assumes that the trait of interest is segregating in the breeding material and hence may not assist in the identification and introgression of novel alleles. Therefore, there will be a continuing requirement for advanced backcross QTL mapping for introgression of novel alleles from wild relatives and a capability for map construction for other special cases. LD mapping strategies will work best where there is strong selection pressure for the trait of interest, so the location and management of field trials and the design and application of laboratory assays is crucial to its success.

Whole genome selection strategies

Much of the emphasis thus far has been examination of single genes or, at least, single traits. Some programs have tackled the issue of simultaneous selection of many genes and the logical extension is to undertake "whole genome" fingerprinting and analysis to enable sophisticated selection strategies to be implemented.

2.1 Linkage block analysis and selection

Breeders have long suspected that certain chromosomal regions carry key clusters of genes ("linkage blocks"), which have been highly conserved during selection. Hayes often uses the term "national parks" to describe these key linkage blocks. The North American Barley Genome Mapping Program (NABGMP) describes regions on chromosomes 1H and 4H associated with a range of malt quality traits. Baum *et al.* (pers. comm.) describe a region on chromosome 5H which contains QTL crucial for superior performance under the stress conditions encountered in Syria. Preliminary evidence from a recent student project at the Waite Campus (Atmodjo, 2000) indicates conserved sections of the genome from the important parental line, CI3576, crucial for adaptation to South Australian conditions, have been retained in current breeding material through 4 or 5 cycles of crossing and selection. Few breeders have attempted to characterise linkage blocks in the breeding material and most do not have the tools in place to exploit this information. Macaulay *et al.*, (2001) describe a set of 48 SSRs, which were used to construct a "genotypic database". This information was used to develop graphical genotypes in MS Excel, an easy method

for visualisation of the conserved regions in a set of breeders' lines. Currently, the most advanced software for "graphic genotypes" resides in the private sector although several programs are currently available in the public domain e.g. "GGT" (Berloo, 1999) and Geneflow[®] (<http://www.geneflowinc.com/>).

2.2 Key recombination events

One of the greatest positives to come from marker development programs has been the increased knowledge of the barley genome and the physical and genetic control of important traits. After the first decade of research in this area, barley breeders can feel much more confident of designing an appropriate strategy to transfer a trait from one parent to another or indeed improve a quantitative trait.

There have been 24 QTL derived from the mapping populations characterised worldwide, associated with the expression of malt extract in barley. Coupled with this information, the location of many important agronomic traits is also known. In the case of breeding malting barley for southern Australia, it is now apparent why it has been difficult to improve malt extract simultaneously with selection for resistance to cereal cyst nematode either at the *Ha2* (2H) or *Ha4* locus (5H). The parental lines used as donors of the resistance carried *Ha2* and /or developmental genes in repulsion to high malt extract QTL from European and Canadian varieties and recombination between the CCN resistance and high malt extract QTL have been rare events in the Australian barley germplasm pool. There are a number of key repulsion phase linkage / pleiotropisms (it is not possible yet to tell which of these two alternatives is the case) which require breeders to develop special strategies are

2HS extract vs. photoperiod

2HL CCN (*Ha2*) vs. photoperiod vs. boron tolerance

5H extract/alpha vs. Kolbach Index

5H extract/alpha vs. post harvest dormancy

3H semi-dwarf vs. small grain vs. late maturity

2.3 Bioinformatics

For whole genome strategies to be easily interpreted and ultimately implemented in practical breeding, analysis and visualisation of the data becomes crucial. In section 2.1, the use of Excel and GGT was briefly discussed. A more powerful tool is Geneflow[®]. This package can manage pedigree, genotypic and phenotypic databases and perform analyses combining this information to give

- graphical genotypes,
- follow allele flow through complex pedigrees,
- compare genotypes from pedigree and backcross programs against the preferred genotype
- and much more! Several examples are shown below (Figure 1 and 2).

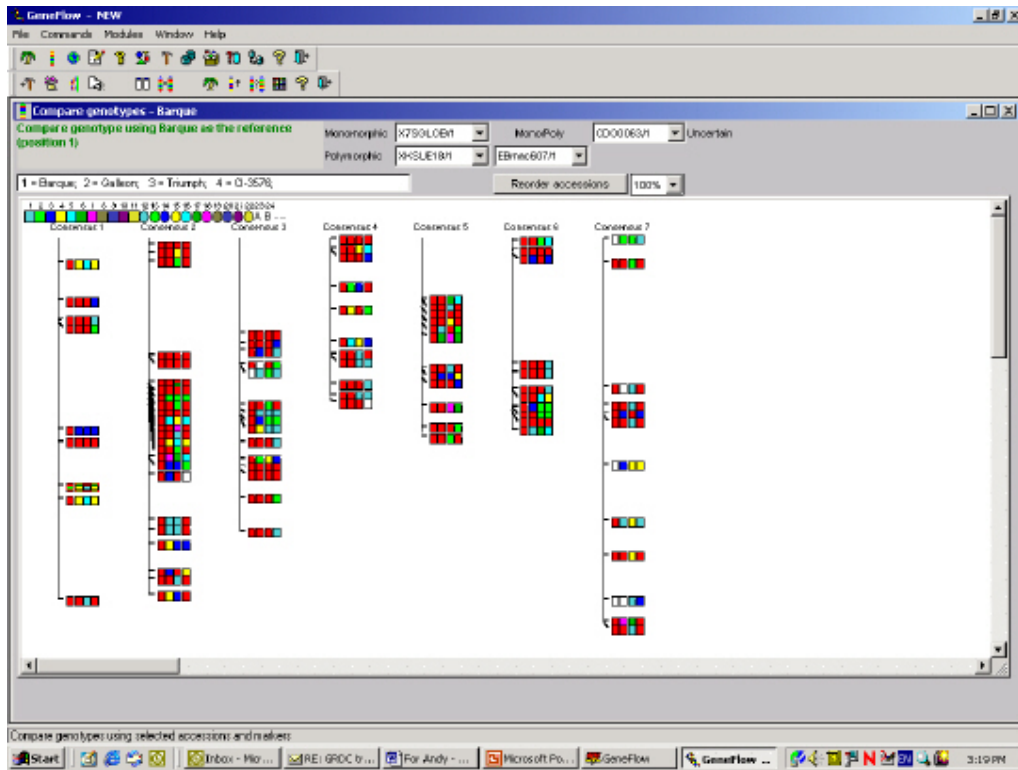


Figure 1 Graphical genotypes prepared in GeneFlow[®] of Barque, Galleon, Triumph and CI 3576 comparing alleles at all available SSR loci - Barque alleles in red

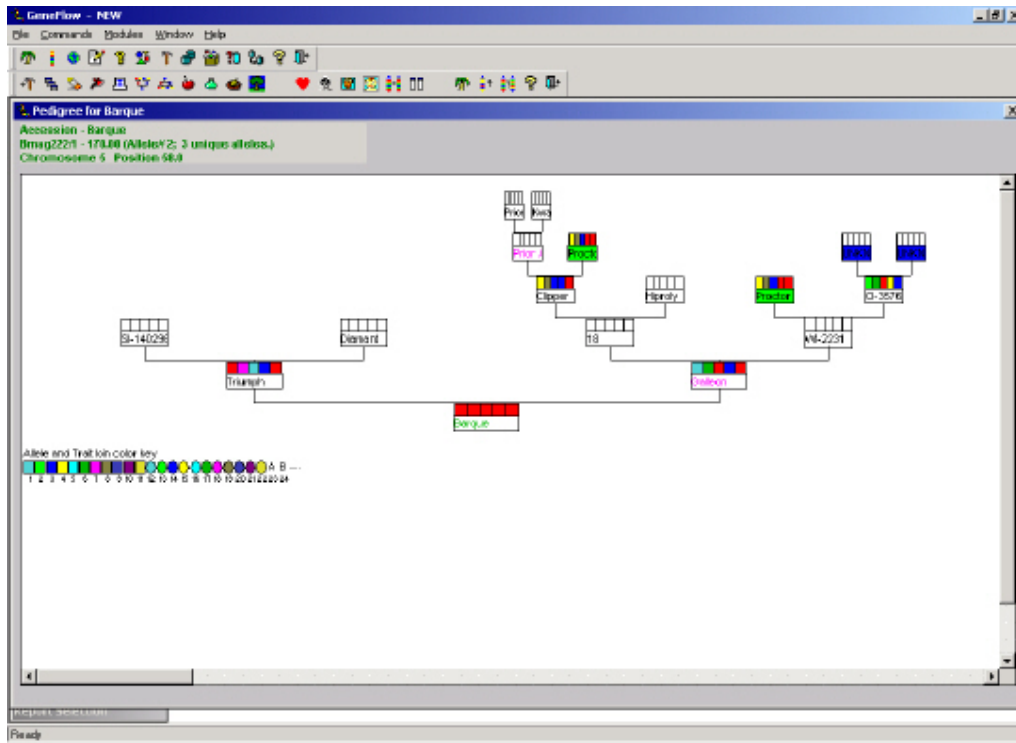


Figure 2 Allele flow in the pedigree of Barque generated in GeneFlow[®] showing origin of the Barque alleles on 2H - Barque alleles shown in red.

2.4 New marker Systems including SNP and DARTs

In addition to the widely implemented SSR markers, newer technologies are also slowly becoming available. Single nucleotide polymorphisms (SNPs) are expected to become an important marker system in a few years time when sequence databases become more extensive. One SNP for β -amylase isoform has recently been developed in WA (Paris, pers. comm.). Particularly encouraging have been microarray based markers systems, such as the Diversity Array (DART) system developed by CAMBIA (Jaccoud *et al.*, 2001). This technique offers the potential to scan a large number of loci simultaneously and is, consequently, particularly promising as a method for fingerprinting lines and varieties.

2.5 Pyramiding malt quality genes

The knowledge of the genome location, size of effect and source of positive QTL for malt extract allows us to pose and answer the obvious question "What happens when we pyramid these QTL?" There are several groups attempting to answer this question in different ways. Some are trying to accumulate QTL from elite parents of mapping populations (such as Morex, Franklin, Alexis, Harrington), others are intercrossing elite individuals from mapping populations, yet others are transferring multiple QTL from one background into a feed variety. The results of these projects may well

determine how useful MAS will be in future efforts to improve the malting quality of Australian barley.

Acknowledgements

The authors gratefully acknowledge the support of the Malting Barley Quality Improvement Program, The GRDC (especially the National Barley Molecular Marker Program) and the CRC for Molecular Plant Breeding. Jodie Kretschmer, Angelo Karakousis and Elysia Vassos who provided skills essential to the conduct the MAS studies at the Waite.

References

1. Atmodjo, M. (2000) Unpub'd Master Ag.Sc. Thesis, Adelaide University, Dept of Plant Science
 2. Barr A.R., S. Roumeliotis, S.J. Logue and S.P. Jefferies (1999) Proc 9th Aust Barley Tech. Symp., Melbourne. Pg 2.45.1 - 2.45.5.
 3. Berloo, R (1999) Genetics 328-329.
 4. Collins, H.M, Logue, S.J, Jefferies, S.P., Stuart I.M., and Barr A.R., (1999) Proc. 9th Australian Barley Technical Symposium, Melbourne, Aust. Pg 2.44.1-2.44.6.
 5. Hayes, P.M. and B.L. Jones (2000) Proc. 8th International Barley Genetics Symposium, Adelaide, 2000 pg 99 - 106.
 6. Jaccoud, D, Peng, K, Feinsein, D and Kilian, A (2001) Diversity arrays: a solid state technology for sequence independent genotyping. Nucl. Acids Res. 29, 25e
 7. Macauley, M., Ramsay, L., Powell, W., and Waugh, R., (2001) A representative, highly informative "genotyping set" of barley SSRs. Theor. Appl. Gen 102: 801 -809.
-