CHAPTER 7

Marker-assisted selection in common beans and cassava

Mathew W. Blair, Martin A. Fregene, Steve E. Beebe and Hernán Ceballos

SUMMARY

Marker-assisted selection (MAS) in common beans (*Phaseolus vulgaris* L.) and cassava (*Manihot esculenta*) is reviewed in relation to the breeding system of each crop and the breeding goals of International Agricultural Research Centres (IARCs) and National Agricultural Research Systems (NARS). The importance of each crop is highlighted and examples of successful use of molecular markers within selection cycles and breeding programmes are given for each. For common beans, examples are given of gene tagging for several traits that are important for bean breeding for tropical environments and aspects considered that contribute to successful application of MAS. Simple traits that are tagged with easy-to-use markers are discussed first as they were the first traits prioritized for breeding at the International Center for Tropical Agriculture (CIAT) and with NARS partners in Central America, Colombia and eastern Africa. The specific genes for MAS selection were the *bgm*-1 gene for bean golden yellow mosaic virus (BGYMV) resistance and the *bc*-3 gene for bean common mosaic virus (BCMV) resistance. MAS was efficient for reducing breeding costs under both circumstances as land and labour savings resulted from eliminating susceptible individuals. The use of markers for other simply inherited traits in marker-assisted backcrossing and introgression across Andean and Mesoamerican gene pools is suggested. The possibility of using MAS for quantitative traits such as low soil phosphorus adaptation is also discussed as are the advantages and disadvantages of MAS in a breeding programme. For cassava, the use of multiple flanking markers for selection of a dominant gene, *CMD2* for cassava mosaic virus (CMV) resistance at CIAT and the International Institute of Tropical Agriculture (IITA) as well as with NARS partners in the United Republic of Tanzania using a participatory plant breeding scheme are reviewed. MAS for the same gene is important during introgression of cassava green mite (CGM) and cassava brown streak (CBS) resistance from a wild relative, *M. esculenta* sub spp. *flabellifolia*. The use of advanced backcrossing with additional wild relatives is proposed as a way to discover genes for high protein content, waxy starch, delayed post-harvest physiological deterioration, and resistance to whiteflies and hornworm. Other potential targets of MAS such as beta carotene and dry matter content as well as lower cyanogenic potential are given. In addition, suggestions are made for the use of molecular markers to estimate average heterozygosity during inbreeding of cassava and for the delineation of heterotic groups within the species. A final section describes the similarities and differences between the MAS schemes presented for the two crops. Differences between the species can be ascribed partially to the breeding and propagation systems of common beans (seed propagated, self-pollinating) and cassava (clonally propagated, cross-pollinating). In addition, differences in growth cycles, breeding methods, availability of genetic markers, access to selection environments and the accompanying opportunities for phenotypic selection influence the decisions in both crops of when and how to apply MAS. Recommendations are made for applying MAS in breeding of both crops including careful prioritization of traits, marker systems, genetic stocks, scaling up, planning of crosses and the balance between MAS and phenotypic selection.

COMMON BEANS: IMPORTANCE AND GENETICS

Common beans (*Phaseolus vulgaris* L.) are the most important grain legume for direct human consumption, especially in Latin America and eastern and southern Africa. They are seed-propagated, true diploids (2n = 22) and have a relatively small genome (650 Mb) (Broughton *et al.*, 2003). Originating in the Neotropics, common beans were domesticated in at least two major centres in Mesoamerica and the Andes (Gepts, 1988) and possibly in a third minor centre in the northern Andes (Islam *et al.*, 2002). Wide DNA polymorphism is expressed between the two major gene pools. Mesoamerican beans typically have small to medium size seeds and can be classed into four races that are distinguished by randomly amplified polymorphic DNA (RAPD) polymorphisms (Beebe *et al.*, 2000). Andean beans usually have medium to large seeds, and landraces have been classed into three races based on plant morphology and agro-ecological adaptation (Singh, Gepts and Debouck, 1991). These can be differentiated by microsatellites (M. Blair, unpublished data) but the genetic distance among Andean races is narrower than that among Mesoamerican races (Beebe *et al.*, 2001). A large number of gene tagging studies have been conducted in common beans, predominantly with RAPD markers, some of which have been converted subsequently to sequence characterized amplified regions (SCARs; reviewed most recently by Miklas *et al.*, 2006).

Beans display a wide range of growth habits (Van Schoonhoven and Pastor-Corrales, 1987), from determinate bush types, to indeterminate upright or viny bush types, to vigorous climbers. Bush types are the most widely grown, and are a relatively short season crop, maturing in as little as 60 days from seeding in a tropical climate and yielding from 700 to 2 000 kg/ha on average. On the other hand, in smallholder agriculture where land is scarce, labour-intensive, high-yielding climbing beans enjoy continuing or even expanding popularity. Climbing beans can mature in 100 to 120 days at mid-elevations, but can delay as long as ten months at higher elevations and can produce the highest yields for the crop, up to 5 000 kg/ha. These features have significant implications for breeding programmes. In bush types it is possible to obtain up to three cycles per year in the field, or even four cycles in greenhouse conditions. Breeding bush beans is thus quite agile with regard to advance of generations, although seed harvest of individual plants is sometimes limited. With climbing beans, on the other hand, at best it is possible to obtain two cycles per year with field grown plants, while managing climbing beans in the greenhouse is logistically difficult. However, while bush beans produce on average 20 to 50 seeds/plant, individual plants of climbing beans often produce enough seeds to plant several rows (100 to 150 seeds).

Beans are self-pollinating and thus breeding methods for autogamous crops are employed. Pedigree selection or some adaptation thereof is most common, and both recurrent (Muñoz *et al.*, 2004) and advanced (or inbred) backcrossing (Sullivan and Bliss, 1983; Buendia *et al.*, 2003; Blair, Iriarte and Beebe, 2003) have been used. Recurrent selection has also been employed (Kelly and Adams, 1987; Beaver *et al.*, 2003) but seldom in a formal sense with a defined population structure. Singh *et al.* (1998) suggested a system that they called gamete selection in which individual F_1 plants of multiple parent crosses give rise

to families. This system takes advantage of the variability among F_1 plants that is created between segregating parental plants. The choice of breeding method and its adaptation to specific circumstances, the growth cycle of the crop in relation to different planting seasons, the access to selection environments and the accompanying opportunities for phenotypic selection and the ease of implementing the specific markers to be used will all influence the decisions about where and how MAS will be most cost effective and used to best advantage.

MAS in bean breeding: experiences of CIAT and NARS

Molecular markers have been sought for both simple and complex traits in beans, with an eye to eventual application in MAS. Tagging of genes and QTL in common bean and their application to MAS have been reviewed previously (Kelly *et al.,* 2003; Miklas *et al.,* 2006). In the present chapter, some of the aspects that contribute to the successful use of MAS are considered in greater detail, referring to examples taken from bean breeding in the tropics at CIAT and within NARS. Simple and complex traits are discussed separately, as they represent two contrasting sorts of experience.

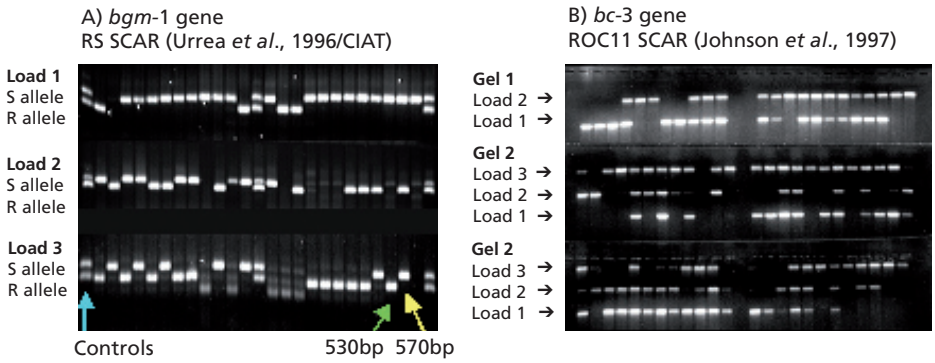
Simple traits

Bean golden yellow mosaic virus resistance

Bean golden yellow mosaic virus (BGYMV) is a white fly-transmitted Gemini virus, and a major production limitation of beans in the mid-to-low altitude areas of Central America, Mexico and the Caribbean. Host resistance to the virus is the most practical means of control, and any new variety in these production areas must carry resistance. Studies on inheritance of resistance revealed a major gene denominated *bgm*-1 in breeding line A429 (Blair and Beaver, 1993) that originates in the Mexican (Durango race) accession "Garrapato" or G2402. Minor genes (Miklas *et al.,* 2000c) as well as additional recessive and dominant resistance genes exist for the virus (Miklas *et al.,* 2006). In most production areas where BGYMV exists, it is necessary to pyramid genes for adequate disease control. Although lines developed in CIAT target these areas, BGYMV does not exist at levels that would permit selection under field conditions in Palmira, Colombia, at CIAT headquarters. Therefore, MAS was desirable to assure recovery of at least the most important resistance genes. MAS has also been employed in the Panamerican School in Zamorano, Honduras, as a complement to field screening, to extend selection to sites and seasons with less disease pressure (J.C. Rosas, personal communication).

A co-dominant RAPD marker was identified for the *bgm*-1 gene (Urrea *et al.,* 1996) that was subsequently converted to a SCAR marker named SR2 (CIAT, 1997). The DNA fragment associated with *bgm*-1 gene has only been observed in one genotype other than G2402 and its derivatives, and thus the polymorphism has been very useful for recognizing the presence of the gene in different genetic backgrounds. This SCAR was evaluated on as many as 7 000 plants in a single sowing (CIAT, 2001; 2003). The uniqueness of the marker's polymorphism and its reliability over laboratories, seasons and genetic backgrounds have facilitated its wide use. More recently, a second SCAR (SW12.700) was developed from the W12.700 RAPD for a QTL located on linkage group b04 (Miklas *et al.,* 2000c), and this has also been incorporated into the breeding programme of CIAT. The combination of *bgm*-1 and the QTL is expected to offer an

FIGURE 1
Examples of gel multiplexing for MAS of A) BGYMV and B) BCMV resistance genes



intermediate level of resistance, while other minor genes must be recovered through conventional phenotypic selection to assure higher resistance.

Scaling up of MAS required the development of simple operational procedures in both the field (tagging, tissue collection) and the laboratory (DNA extraction, marker evaluation). For gamete selection strategies in the field, individual, evenly-spaced plants from segregating populations were marked with numbered tags that were coated with paraffin to protect them until seed harvest. Leaf disks were sampled from young vegetative tissue with a paper hole puncher and placed directly into pre-numbered cells of microtitre 96-well plates stored on ice, ready for grinding and extraction in the laboratory. The implementation of MAS for *bgm*-1 and subsequently for SW12.700 in the laboratory required substantial adaptation of standard protocols to establish high-throughput procedures. Grinding of samples in microtitre plates was accomplished with a block of 96 pegs that fit into each well. Alkaline DNA extraction (Klimyuk *et al*., 1993) was employed with success for both markers, and eventually it was possible to multiplex the markers in both the amplification and gel phases using multiple primer PCR and multiple loading per gel wells (Figure 1A). With experience and improved procedures, efficiency more than doubled over a two-year period. MAS was often carried out before flowering to decide on a plant's status as a carrier of the resistant allele for further use in crossing.

Two small red seeded lines developed in the Panamerican School using MAS have reached the stage of validation in Honduras (J.C. Rosas, personal communication) and shown resistance to the BGYMV strains prevalent there. Resistance to BGYMV of drought tolerant lines selected at CIAT was maintained using MAS for one or more genes, followed by field selection in Central America. Similarly, red mottled lines developed in CIAT with the aid of MAS showed field resistance in the Caribbean and one of these lines from the red mottled advanced line for the Caribbean (RMC) series has been released (Blair *et al.,* 2006). MAS has also been an important element of maintaining BGYMV viral resistance in CIAT's programme as other breeding objectives such as nutritional value have been assumed, necessitating the inclusion of susceptible parents in crosses with resistant

lines. MAS for this trait has also been practised at the University of Puerto Rico and at the Biotechnology Institute of Cuba.

Bean common mosaic virus and bean common mosaic necrotic virus

Bean common mosaic virus (BCMV) and the related necrotic strains (bean common mosaic necrotic virus [BCMNV]) are aphid-transmitted potyviruses that are found worldwide and are seed-borne from season to season. BCMNV resistance is very important in Africa where necrotic strains are prevalent and has become a renewed priority for parts of the Caribbean where necrotic strains have been discovered. BCMV is also endemic in the Andean region where it persists in farmer-saved seed and long-season climbing beans. Climbing beans are grown in both intensive (trellised/staked monoculture) and extensive (intercropping with maize) farming systems. In both systems the need to protect the crop from easily transmitted viral diseases such as BCMV or BCMNV is great; however, very few climbing beans have been bred for resistance to BCMV. A number of BCMV/BCMNV resistance genes have been tagged including the dominant *I* gene (with which the necrotic strains interact to produce necrosis) and the recessive *bc*-3, *bc*-2 and *bc*-1^2 genes (Haley, Afanador and Kelly, 1994; Melotto, Afanador and Kelly, 1996; Johnson *et al.,* 1997; Miklas *et al*., 2000a). The genes can be distinguished by inoculation with different viral isolates, and a range of molecular marker tags are available for each gene (reviewed in Kelly *et al*., 2003; Miklas *et al.,* 2006). The dominant *I* gene was incorporated into a wide range of small seeded bush beans at CIAT, while resistant bush beans of the bush bean resistant to black root (BRB) series carrying recessive genes were developed in the 1990s and have been widely distributed as breeding parents. The need to reselect the recessive genes with confidence from segregating populations makes MAS a priority.

CIAT started a collaborative project with the Colombian national bean programme based at the Colombian Agricultural Research Corporation (CORPOICA) in 2002 to introgress BCMV resistance genes from BRB lines into local landraces and improved genotypes of Andean climbing beans (CIAT, 2002, 2003, 2004; Santana *et al.,* 2004). During the breeding programme for BCMV and over the course of four years, MAS was used extensively based primarily on the SCAR marker ROC11 developed for the *bc*-3 gene (Johnson *et al.,* 1997) and the SCAR marker SW13 for the *I* gene (Melotto, Afanador and Kelly, 1996) along with virus screening to confirm the selection of resistant progeny. The programme was successful in moving *bc*-3 resistance into a background of cream mottled and red mottled seed types for both highland areas (known as Cargamanto commercial class) as well as mid-altitude areas through triple-, double- and backcrosses. Although virus resistance was also screened phenotypically, the frequency of escape, the complex interaction of multiple genes and the recessive nature of most of these made MAS the best option for breeding resistant varieties rapidly. In addition, as climbing bean breeding is a more time-consuming and expensive endeavour than bush bean breeding due to the longer season, wider plant spacing and need for staking material, MAS was also found to be a very effective measure to reduce breeding costs and save on breeding nursery space.

The implementation of MAS for BCMV was based on a combination of the previously developed SCAR markers previously mentioned and techniques developed at

CIAT for the selection of BGYMV resistance as discussed previously. Although most BCMV and BCMNV resistance genes had been tagged with SCAR markers, implementation required efforts to validate and scale up the use of the markers in applied breeding programmes. Genotyping for the ROC11 marker was carried out on advanced lines given that this marker is dominant and in repulsion with the resistance allele. In other words, the absence of a band was indicative of the presence of the recessive *bc*-3 allele and therefore it was more appropriate to evaluate after fixation of the alleles to homozygosity through mass or pedigree selection with single plant selections in the F_4 and F_5 generation when single plant rows were evaluated for the resistance gene marker. To determine whether the advanced line continued to segregate for the gene, alkaline DNA extraction was conducted on leaf discs collected from four leaflets from four individual plants per line using a hole-puncher rather than from a single plant per family or advanced line. The presence or absence of polymerase chain reaction (PCR) products was evaluated for each genotype based on scanned photographs or gel capture imagery of multiplexed gels (Figure 1B) to predict if the genotype contained the resistance or the susceptible allele.

Once optimized for parental genotypes, MAS was conducted on a large number of progeny rows. For example in 2003, more than 4 000 advanced lines were evaluated for the ROC11 marker for genotypes grown at three sites within Colombia (CIAT-Darien, CIAT headquarters and CORPOICA-Rionegro). DNA was collected at all three sites and shipped successfully to the laboratory in 96-well plate format as discussed above. Both the ROC11 and SW13 markers were single copy SCARs that did not produce extra bands and therefore were easy to multiplex. To facilitate the evaluation of markers on a large number of advanced lines, usually within two to three weeks, and increase the efficiency of MAS, several innovations were implemented: loading of agarose gels (first with two and then three loadings), increasing numbers of wells per comb (first 30-well and then 42-well combs were used), use of 384-well PCR plates and multipipetor loading of gels. The resulting savings decreased the time to PCR amplify and load a gel by approximately 50 percent and increased the number of genotypes run per gel by 225 percent.

The rapid increase in efficiency obtained during the application of the ROC11 marker shows the advantages of testing new markers in practical breeding programmes. The use and advantages of these molecular markers has been presented at an Organization of American States-sponsored course in Colombia given in 2002 and a Rockefeller Foundation-sponsored course in Uganda given in 2003. Based on this programme and the training courses, MAS for BCMV genes was initiated as part of a recently approved Association for Strengthening Agricultural Research in Eastern and Central Africa (ASARECA) project for three countries in eastern Africa and training of researchers from the Andean region has allowed more breeding lines from Peru to be screened (CIAT, 2004).

Other examples of MAS for simply inherited traits

Several pathogens, especially fungal pathogens, have co-evolved with the bean host, and present a population structure (Andean/MesoAmerican) that mimics the major gene pools of bean (Pastor-Corrales, Jara and Singh, 1998). This is the case with *Phaeoisariopsis griseola*, the

causal agent of angular leaf spot (ALS), and *Colletotrichum lindemutheanum*, which induces anthracnose. In both cases, pathogen isolates tend to be more virulent on host genotypes of the same gene pool (Andean or Mesoamerican) and less so on host genotypes from the contrary gene pool. Resistance genes of utility to one host gene pool thus tend to originate in the other gene pool and require introgression from one gene pool to the other. MAS has great potential for introgression as DNA polymorphisms are maximized in wide crosses across gene pools, and markers are available for this purpose for both ALS (Carvalho *et al.,* 1998; Sartorato *et al.,* 1999; Nietsche *et al*., 2000; Ferreira *et al.,* 2000; Mahuku *et al.,* 2004) and anthracnose (Young *et al*., 1998; Awale and Kelly 2001; Vallejo and Kelly, 2001).

Other cases of wide crosses in which MAS can be of use include those for the selection of genes for resistance to a storage insect, the Mexican bean weevil (*Zabrotes subfasciatus* [Boheman]) derived from wild bean accessions from Mexico. Selection for resistance has also been achieved by analysis for the active resistance agent, a seed protein called arcelin, by either antibody reaction or electrophoresis, but MAS is simpler and more efficient than either of these analyses that require protein extraction. Even wider crosses of common bean with *Phaseolus acutifolius* have recovered resistance to common bacterial blight (caused by *Xanthomonas axonopodis* pv. *phaseoli*) (Muñoz *et al.,* 2004) and markers have also been developed for these resistance genes (Jung *et al.,* 1997; Miklas *et al.,* 2000b; Park *et al*., 1999; CIAT, unpublished data). In these cases also, the fact of deploying genes from relatively wide crosses favours maintaining a state of DNA polymorphism in relation to the target genotypes.

Complex multigenic traits

In addition to the studies previously discussed, several attempts have been carried out to tag quantitative trait loci (QTL) for abiotic stress tolerance or insect resistance in common bean, although most of these traits might better be described as oligogenic, as results usually suggest that a limited number of loci (from three to six) are involved in their genetic control.

One example is tolerance to low soil phosphorus that was investigated in the landrace G21212. Linkage group b08 proved to be especially important to yield under low phosphorus, with as many as three important and loosely linked QTL (Beebe, Velasco and Pedraza, 1999; Miklas *et al.,* 2006). Interestingly, these same QTL were linked to QTL for resistance to *Thrips palmi* Karny derived from the same source (Frei *et al.,* 2005). This is a promising candidate for applying MAS in the short term for abiotic stress tolerance, although another notable attempt was also made for drought tolerance breeding with MAS through a joint programme between Michigan State University and the National Institute for Forestry, Agriculture and Livestock Research (INIFAP) in Mexico (Schneider, Brothers and Kelly, 1997).

In theory, a breeder would prefer markers for low heritability quantitative traits that are difficult to select through phenotypic selection. However, in general, markers for polygenic or oligogenic traits have not moved into the application phase. The same problems that make phenotypic selection difficult apply in some degree to MAS. Multiple minor genes that are often associated with poor heritability also imply that it is difficult to identify QTL with highly significant effects and that merit the investment of MAS. Furthermore, good genome coverage is usually necessary to

detect the QTL that explain the highest amount of genetic variability, and this has been difficult to achieve in intragene pool crosses in common beans.

However, genetic analysis by markers has been very useful for revealing the inheritance of quantitative traits, especially physiological traits, even when the markers involved did not result in application in MAS. Analysis of QTL was applied to root traits of bean as they relate to absorption of phosphorus from soil (Liao *et al.*, 2004; Yan *et al.*, 2005; Beebe *et al.*, 2006). This permitted associating different physiological traits to P uptake and estimating their importance in nutrient acquisition. Once traits are better understood, then an appropriate selection strategy can be devised, be it phenotypic or MAS. Thus, markers can be useful to a breeding programme by elucidating basic plant mechanisms even if they are not applied directly in selection.

Breeding schemes: adaptation to include MAS

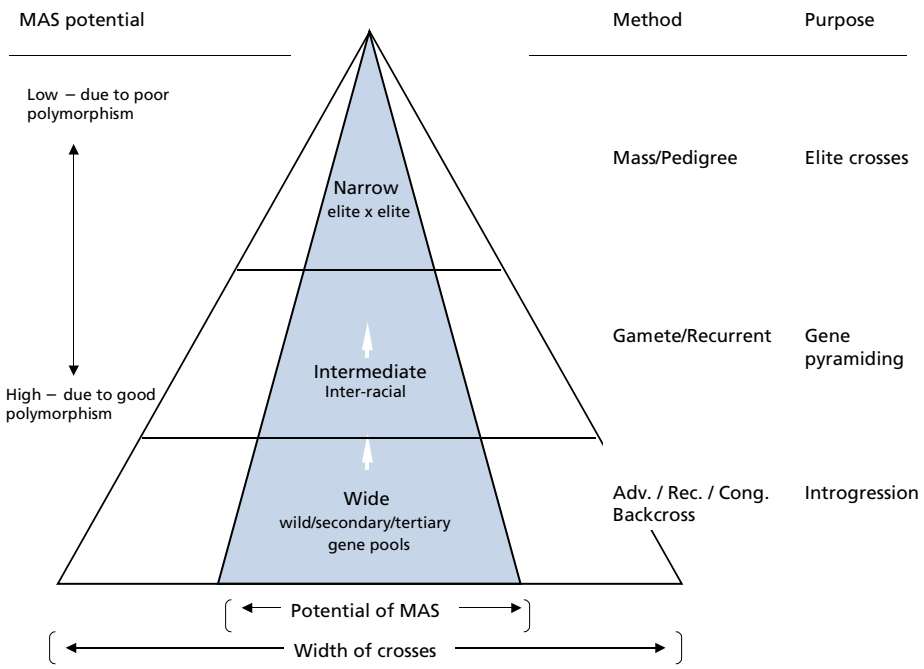
The eventual application of MAS requires careful prioritization of traits and even specific genes for which markers are to be sought, in light of the importance of the trait and genes, and options for phenotypic selection. One should never assume that MAS is necessarily superior to phenotypic selection, which for some traits may be as effective and efficient as the use of molecular markers. However, if a gene is sufficiently important in a breeding programme to demand that advanced lines have such a gene (as in the case of the *bgm*-1 gene for virus resistance in Central America), there is probably some point in the selection process at which MAS would be useful. Also, it is not necessary to select many genes by MAS for it to be of great value. For example, if a single gene is segregating and 50 percent of plants lack the gene in advanced generations, an effective selection would eliminate half the population and increase the subsequent efficiency of the breeding programme by a factor of 2.

Once markers are available, a key issue is determining the range of parental genotypes within which a marker is polymorphic and therefore useful for selection. Markers of genes that originate from wider crosses (e.g. from different races, gene pools or species) will have a progressively greater chance of being polymorphic among a range of parents (Figure 2) and therefore diagnostic for the gene of interest. The example of *bgm*-1 is again a good case in point as the resistance allele and the SR2 marker are both unique to the Durango gene pool and polymorphic in combinations across other Mesoamerican races as well as the Andean gene pool. In contrast, the ROC11 marker for the *bc*-3 gene is only polymorphic across gene pools and therefore not diagnostic for the resistant allele.

If a breeder has several potential parents among which to choose and these are comparable with regard to other traits, it might be preferable to eliminate those that carry a band that would be confused with the linked marker and would result in false positives. Conversely, if more than one marker is available for a given gene, one might focus on those linked markers that maintain polymorphism in the greater number of combinations. In some combinations it might be informative to use both linked markers simultaneously, both to discern recombinants and to confirm markers.

Several possible schemes for the introduction of MAS to different breeding schemes are represented in Figure 3. A breeder must consider at what generation in the breeding programme selection

FIGURE 2
Potential of MAS in crosses of varying genetic diversity



Modified from Kelly, Schneider & Kolkman, 1999 and Singh, 1999

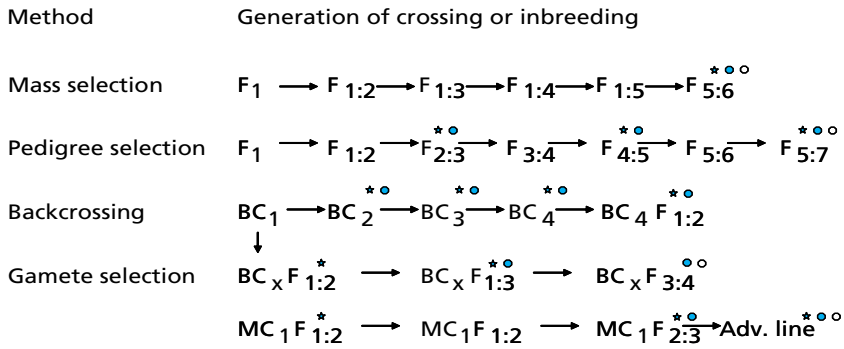
by MAS will give the greatest cost/benefit ratio. This would probably be early in the breeding programme for the pedigree method or for gamete selection while it would be later in the programme for bulk method or mass selection (Figure 3). In the case of early generation selection, elimination of plants without the gene(s) will avoid unproductive investment.

Advantages and disadvantages of MAS

MAS provides real advantages where the conditions are not favourable for phenotypic selection, for example, in the case of BGYMV, which does not exist at epiphytotic levels in CIAT. Indeed, *bgm*-1 behaves as a recessive gene, so phenotypic selection in early generations would be inefficient in recovering the gene in the heterozygous state.

The same principle would apply to the recessive *bc*-3 gene, although the lack of a marker linked in coupling to this gene has been a serious drawback and has limited the effectiveness of MAS to advanced generations when the gene is fixed by inbreeding. In this case, early generation selection with MAS would be limited to negative selection against homozygous dominant and heterozygous plants, and this eliminates potentially useful allele-carrying genotypes. Indeed, MAS is impossible in generations such as the F_1 or BC_1F_1 to the susceptible parent when no homozygous recessive plants exist at all.

FIGURE 3
Application of co-dominant and dominant markers during the breeding cycle in common beans



Note: Asterisks indicate co-dominant marker, while open circles indicate dominant marker in repulsion and closed circles indicate dominant marker in coupling.

In other cases where phenotypic selection methods are available, the advantage of MAS resides in its simplicity. This is the case in the selection of arcelin, which can be achieved through protein extraction followed by antibody detection or electrophoresis, but both of these are laborious while MAS can be applied more rapidly and with much greater throughput. Similarly, markers for common blight resistance and anthracnose have the advantage of obviating the need for field inoculations that are sometimes ineffective if environmental conditions are not favourable. The advantage of MAS is much greater if a single DNA extraction can serve for the evaluation of several markers, as in the multiplexing of *bgm*-1 and SW12.700 markers.

In spite of attempts to apply MAS to complex traits, examples of successful application are still limited to relatively simple traits. This is contrary to some previous expectations that markers would benefit mostly traits of low heritability. However, experience has shown that the ability to manipulate even one important gene with confidence can make a breeding programme more efficient, if that gene is highly desirable and valuable for advanced materials.

Meanwhile the disadvantages of MAS compared with phenotypic selection are based on effectiveness and cost considerations. The effectiveness of MAS is relative to the ease of applying a given marker, its reliability and its level of linkage with the gene of interest. Although molecular markers theoretically have a heritability of 1.0, variability among laboratories or among runs within a laboratory make markers less than 100 percent reliable. This is especially true for RAPD markers for which band amplification is dependent on DNA concentration and quality, annealing temperature and thermocycling conditions, Taq polymerase concentration and the relative proportion of various other ingredients to the PCR cocktail. In comparison, SCAR markers are much more reliable and repeatable and therefore have higher heritability than RAPD markers. Linkage distance between a marker for a

gene of interest and the actual locus itself also affects the reliability of a marker. In turn, the type of cross (wide versus narrow) and parents involved (closely or distantly related) affect the frequency of recombination around introgressed genes as well as the level of polymorphism of the cross and whether the marker will present distinct alleles for the desirable and undesirable character states. In this regard, there is a tradeoff as MAS is most effective when there is high polymorphism in the crosses being evaluated (Figure 2). However, this is precisely the breeding situation in which gene introgression is most difficult, time-consuming and plagued by linkage drag, as is the case for interspecific or interspecific-derived crosses, hybridizations with wild or wild-derived genotypes and crosses between the Andean and Mesoamerican gene pools. This issue is being addressed in beans with the development and mapping of microsatellite markers (Blair *et al.,* 2003) that are much more polymorphic and useful for diagnosing the inheritance of genomic segments in narrow crosses. The first application of microsatellite markers for MAS in common beans was the selection of arcelin based bruchid resistance using gene-derived simple sequence repeats that are diagnostic for the introgression of alleles for resistance from wild beans into cultivated backgrounds (CIAT, 2004), but others should also show promise.

In terms of cost considerations, the relative costs of MAS versus phenotypic selection are relative to each trait and situation. The widely held perception that MAS is expensive is often due to the ingredients and time used to prepare DNA extractions and PCR reactions, although these costs have been reduced by innovations such as the alkaline DNA extraction technique (Klimyuk *et al.,* 1993) that obviates the need for organic solvents or expensive enzymes involved in other mini-preparation techniques (Afanador and Hadley, 1993). While experienced labour was previously required for DNA extraction at CIAT or in NARS breeding programmes, the alkaline extraction method allows most laboratory steps to be carried out even by untrained personnel. Furthermore, MAS costs can be reduced by miniaturization, especially in the PCR reaction (for example, use of 384-well PCR plates and small reaction volumes) and re-use of ingredients (for example plasticware including pipette tips and microtitre plates as well as agarose from used gels). As previously mentioned, multiplexing adds to the efficiency and therefore reduces the datapoint costs of MAS.

Currently, MAS with SCAR markers and alkaline extraction at CIAT cost less than US$0.25 per datapoint. Therefore the expense of MAS is now not as important an issue as previously. In this regard, MAS sometimes has the advantage of being implemented in any generation and under both field or greenhouse conditions, while phenotypic selection often requires a separate planting and specialized labour for inoculation, agronomic management and evaluations or scoring. However, in the final analysis, the most efficient and cost effective breeding programme will probably be one that combines MAS and phenotypic selection in some optimal combination. It is precisely the challenge of the breeder to define that optimal combination.

One last disadvantage of relying on MAS is that it commits a breeder to a unique gene(s) for a given trait. For example, there might be multiple genes or gene combinations for resistance to a disease, or for a physiological trait such as root structure. To the extent that a breeder relies on MAS for selection, this excludes other

possible genes and the use of other potentially useful parents that do not share the DNA polymorphism that is used in MAS. On the other hand, phenotypic selection would permit recognizing different genetic options for a desired phenotype. Thus, MAS is most useful when it is applied to truly unique genes.

CASSAVA: IMPORTANCE AND GENETICS

Cassava is a perennial shrub but it is generally harvested as an annual crop at 10–11 months of age. Basically every part of the plant can be utilized. The starchy roots are a valuable source of energy and can be boiled or processed in different ways for human consumption and different industrial purposes such as starches, animal feed or alcohol (Ceballos *et al*., 2006). Cassava storage roots are not tubers and therefore cannot be used for reproductive purposes; stems are the common planting materials. Cassava foliage is not widely exploited in spite of its high nutritive value (Buitrago, 1990; Babu and Chatterjee, 1999). Foliage consumption by humans is relatively common in certain countries of Africa, Asia and Latin America. The use of foliage for animal feeding is generating increased interest in Asia.

Cassava can be propagated by either stem cuttings or botanical seed. However, the former is the practice most widely used by farmers for multiplication and planting purposes. Propagation from true seed occurs under natural conditions and is common in breeding programmes. Occasionally botanical seed is also used in commercial propagation schemes (Rajendran *et al.,* 2000).

Cassava is monoecious and allogamous, with female flowers opening 10–14 days before the male ones on the same branch. Pollination can be done manually in a controlled way to produce full-sib families or else in polycross nurseries where open pollination takes place and, therefore, half-sib families are produced. Self-pollination is feasible when using male and female flowers on different branches or on different plants of the same genotypes (Jennings and Iglesias, 2002). Some clones flower relatively early at four or five months after planting whereas others only do so at eight to ten months after planting. As a result, the time required for the seed to mature, the growing cycle of the crop and the need to plant with the arrival of the rains take about two years between a given cross being planned and the respective seed becoming available. On average, between one and two seeds (out of the three possible in the trilocular fruit) per pollination are obtained (Kawano, 1980; Jennings and Iglesias, 2002).

Breeding objectives

Productivity plays a major role in industrial uses of cassava, whereas stability of production is fundamental in the many regions where cassava is the main subsistence crop. Industrial uses of cassava require high dry matter content as the main quality trait for the roots, whereas for human consumption the emphasis is on cooking quality, frequently even over productivity, as the determining trait. Stability of production is associated with resistance or tolerance to major biotic and abiotic stresses, with the emphasis varying with the target environment. Genetic resistance to the most important diseases and pests and the prevalent abiotic stresses can be found in cassava germplasm (Hillocks and Wydra, 2002; Bellotti *et al.,* 2002; Belloti, 2002; Ceballos *et al.,* 2004). Although cyanogenic glucosides are found in every tissue except

the cassava seed, most processing methods allow a rapid release and elimination of the cyanide. Depending on the end use, high or low cyanide clones are preferred. Other relevant traits for the roots are dry matter, protein and carotenoid content (Chávez *et al.*, 2005).

Breeding schemes

Genetic improvement of clonally propagated non-inbred crops such as cassava is made possible by the fact that a superior genotype can be fixed at any stage in the breeding scheme, even after a single cross, the equivalent of an F_1 in commercial hybrids such as maize. Therefore, non-additive gene actions including dominance and epitasis become important components of the genetic variance to be manipulated by the breeder (Jaramillo *et al.*, 2005; Calle *et al.*, 2005; Perez *et al*., 2005a). Large effective breeding population sizes are required to retain favourable dominant alleles and epistatic loci combination.

As in most crop breeding activities, cassava genetic improvement starts with the production of new recombinant genotypes derived from selected elite clones. Scientific cassava breeding began only a few decades ago, and the divergence between landraces and improved germplasm is not as wide as in other crops. Therefore, accessions for germplasm bank collections from different research institutions play a more relevant role in cassava than in other crops that have been scientifically bred for longer periods of time. Parental lines are selected based mainly on their performance *per se* and little progress has been made to use general combining ability (Hallauer and Miranda Fo, 1988) as a criterion for parental selection. Sexual seeds obtained by the different crossing schemes are germinated to initiate a new cycle of selection. The multiplication rate of cassava planting material is low as five to ten cuttings can be obtained from one plant. This implies a lengthy selection process, and in fact it takes about six years from the time the botanical seed is germinated until enough planting material is available for multilocation replicated trials.

Table 1 illustrates a typical selection cycle in cassava. It begins with the crossing of elite clones and finishes when the few clones surviving the selection process reach the stage of regional trials across several locations. It should be emphasized that there is some variation among the few cassava-breeding programmes in the world with respect to the number of genotypes

TABLE 1
Typical selection cycle in cassava beginning with the crossing of elite clones to the point when few clones surviving the selection process reach the stage of regional trials across several locations

Year	Activity	Number	Plants per genotype
1-2	Crosses among elite clones planned, nurseries planted and pollinations made	Up to 100 000	
3	F_1: Evaluation of seedlings from botanical seeds. Strong selection for African cassava mosaic virus (ACMV) in Africa.	100 000[a]; 50 0000[b]; 50 000 [c]	1
4	Clonal evaluation trial (CET)	20 000–30 000 [a, b] 700 [c]	6–8 (1 rep, 1 location)
5	Preliminary yield trial (PYT)	100 [a]; 300 [b]; 80 [c]	20–60 (3 reps, 1 location)
6	Advanced yield trial (AYT)	25 [a]; 100 [b]; 20–25 [c]	100–500 (3 reps, 2–3 location)
7-9	Regional trials (RT)	5-30 [a, b, c]	500-4 000 (3 reps, 3–4 locations)

Figures for cassava breeding at [a] IITA (Ibadan, Nigeria); [b] CIAT (Cali, Colombia) and [c] CIAT and Rayong Field Research Station from Department of Agriculture (Thailand).
Source: adapted from Jennings and Iglesias, 2002.

and plants representing them through the different stages. Table 1 also provides an idea of the selection pressures generally applied.

Strong emphasis on highly heritable traits (plant type, branching habits and reaction to diseases, harvest index and dry matter content) is applied during the early phases of selection (F_1 and CET), (Hahn, Howland and Terry, 1980; Hahn, Terry and Leuschner, 1980; Hershey, 1984; Kawano, 2003; Ceballos *et al*., 2004). As the number of plants representing each genotype increases, the weight of selection criteria shifts towards low heritability traits such as root yield. The clones that show outstanding performance in the regional trials are released as new varieties and, eventually, incorporated as parents in the crossing nurseries. With that the selection cycle is finished and a new one begins. The whole process has the following characteristics (Ceballos *et al.*, 2004):

- the process is indeed phenotypic selection because no family data are involved;
- no data are collected in the early stages of selection. Therefore, data regarding general combining ability effects (~ breeding value) are not available for a better selection of parental materials;
- there is no proper separation between general (GCA ~ additive) and specific (SCA ~ heterotic) combining ability effects. The outstanding performance of selected materials is likely to depend on positive heterotic effects that cannot be transferred to the progenies that are sexually derived from them;
- no inbreeding is incorporated purposely in the selection process. Therefore, large genetic loads are likely to remain hidden in cassava populations and useful recessive traits are difficult to detect;
- several stages of selection are based on unreplicated trials. A large proportion of genotypes is eliminated without proper evaluation.

For the above-mentioned reasons, cassava breeding is difficult, expensive and to a certain degree inefficient (Perez *et al.*, 2005a; Cach *et al*., 2005a, b). Kawano *et al.* (1998) mention that, during a 14-year period about 372 000 genotypes derived from 4 130 crosses were evaluated at the CIAT-Rayong Field Crop Research Center. Only three genotypes emerged from the selection process to be released as official varieties. Similar experiences have been observed at the International Institute of Tropical Agriculture (IITA), CIAT-Colombia and Brazil. Therefore, the development and adaptation of molecular tools for cassava genetic improvement offer important advantages to make the process more efficient and effective.

MAS in cassava breeding

Cassava genetic improvement can be made more efficient through the use of easily assayable molecular genetic or DNA markers (MAS) that enable the precise identification of genotype without the confounding effect of the environment, thereby increasing heritability. MAS can also contribute to the efficient reduction of large breeding populations at the seedling stage based upon "minimum selection criteria". This is particularly important given the length of the growing cycle of cassava and the expense involved in the evaluation process. Therefore, a pre-selection at the F_1 phase (see Table 1) could greatly enhance the efficiency of the CET experiments. The selection of progenies based on genetic values derived from molecular marker data substantially increases the rate of genetic gain, especially if the number of cycles of evaluation or generations can be reduced (Meuwissen, Hayes and Goddard, 2001).

Another application of MAS in cassava breeding is reducing the length of time required for the introgression of traits from wild relatives. Wild relatives are important sources of genes for pest and disease resistance in cassava (Hahn, Howland and Terry, 1980; Hahn, Terry and Leuschner, 1980; Chavarriaga *et al.,* 2004), but the need to reduce or eliminate undesirable donor genome content and linkage drag can lengthen the process, making it unrealistic for most breeders. Simulations by Stam and Zeven (1981) indicate that markers could reduce linkage drag and would reduce the number of generations required in the backcross scheme. Hospital, Chevalet and Mulsant (1992) corroborated this in achieving a reduction of two backcross generations with the use of molecular marker selection. Frisch, Bohn and Melchinger (1999), through a simulation study, found that use of molecular markers for the introgression of a single target allele saved two to four backcross generations. They inferred that MAS had the potential to reach the same level of recurrent parent genome in generation BC_3 as reached in BC_7 without molecular markers.

The decision to employ DNA-based markers in cassava breeding is primarily based on the heritability of a trait and the amount of genotypic variance explained by the marker. There are many instances in cassava breeding where h^2 is low or zero. Some examples are:

- plant health traits where the pathogen or pest pressure is absent or low, such as cassava mosaic disease (CMD) in the New World tropics or cassava green mite (CGM) during the wet season;
- variable or erratic pest pressure, e.g. the CGM or diseases such as the cassava frog skin disease (FSD);
- evaluation based upon a single plant;
- variable experimental fields and/or poor management resulting in large experimental errors;
- traits that are affected by the stage of plant growth or the part of the organ used for tissue analysis, e.g. cyanogenic potential.

In the above-mentioned instances, having a marker(s) that explains a large proportion of the genetic variance can accelerate progress in breeding. Even where h^2 is moderate or high, selection by markers can be advantageous:

- where different sources of genes exist for the trait that are indistinguishable by phenotype alone and pyramiding is difficult and time consuming, e.g. for different sources of resistance to a disease or pest;
- where molecular tags that can be used inexpensively and rapidly to identify desirable genotypes early in the breeding cycle exist, thereby eliminating the need to evaluate large numbers of plants phenotypically, and obviating the confounding effects of the environment. Markers may permit the efficient elimination of undesirable genotypes at the seedling stage. For example, the number of genotypes at the seedling stage can be reduced by 50 percent if a trait is controlled by a single gene, or by 87.5 percent if controlled by three genes;
- for the introgression of useful genes from exotic germplasm into adapted gene pools. MAS can be used to identify genotypes that carry minimal amounts of flanking donor parent genome around the gene of interest for faster backcrossing;
- for definition of heterotic pools in a group of germplasm accessions for more directed crosses;

- for definition of average heterozygosity in the selection of partially inbred lines for tolerance to inbreeding;
- for identification of the male parent in elite germplasm derived from polycrosses by fingerprinting. This tool is also useful for checking the identity of different genotypes to eliminate duplication in germplasm collections.

Best results are achieved when MAS is combined with phenotypic data as compared with either approach independently (Hospital, Chevalet and Mulsant, 1992). Phenotypic data would reduce the cost of genotyping especially if phenotypic evaluation is conducted on early generations (Gimelfarb and Lande, 1994). This not only reduces the cost of MAS but also increases its efficiency. Some examples of MAS in cassava breeding conducted at an international centre and national programmes are described below.

Molecular MAS for CMD resistance at an IARC

An ideal target for MAS is breeding for disease resistance in the absence of the pathogen. This is the case of CMD in the Americas, where the disease does not occur. CMD is a viral disease first reported by Warburg in 1894 in eastern Africa (quoted by Storey and Nichols, 1938). Several variants of the disease (East Africa cassava mosaic virus [EACMV], South Africa cassava mosaic virus [SACMV], Indian cassava mosaic virus [ICMV]) have been reported (Swanson and Harrison, 1994) and are endemic in all cassava growing regions of Africa and southern India, where it is the most severe production constraint. The white fly vector of CMD, *Bemisia tabaci* biotype A, does not colonize cassava in the New World but recently a new biotype of *B. tabaci*, biotype B (also referred to as *B. argentifolia*), has become widespread in the Americas and has a wide host range including cassava (Polston and Anderson, 1997), increasing the possibility that CMD, EACMV, SACMV, ICMV or a native American gemini virus will become established on cassava in the neo-tropics. This is a frightening prospect for cassava production in Latin America, considering that most Latin American cassava germplasm is very susceptible to CMD (Okogbenin *et al*., 1998). The susceptibility of neo-tropical germplasm to CMD also limits the utilization of germplasm from the crop's centre of diversity in the neo-tropics for these key cassava production regions. Breeding for resistance to CMD in Latin America, where the disease does not exist and is unlikely to be introduced due to very strict quarantine controls, requires the tools of MAS.

Evaluations at IITA identified an excellent source of resistance to CMD in some Nigerian landraces (A.G.O. Dixon 1989, unpublished data), namely TME3, TME7, TME5, TME8, TME14 and TME28. This resistance is effective against all known strains of the virus, including the virulent Ugandan variant (UgV) (Akano *et al.,* 2002; CIAT, 2001). CIAT, in collaboration with IITA in Ibadan, Nigeria, and with support from the Rockefeller Foundation, devel-oped several molecular markers for this source of CMD resistance, revealed to be controlled by a single dominant gene designated as *CMD2* (Akano *et al.,* 2002). At least five markers tightly associated to *CMD2* have been developed, the closest being RME1 and NS158 at distances of four and seven cM respectively. The dominant nature of *CMD2* and its effectiveness against a wide spectrum of viral strains makes its deployment very appealing for protecting cassava against the actual or potential ravages of CMD

in both Africa and Latin America. CIAT and IITA undertook a project to verify the utility of these markers for MAS in breeding CMD resistance by developing crosses between the sources of TME3 and susceptible varieties. A total of six families, ranging in size from 36–840 genotypes, and a total of 2 490 genotypes were used. The crosses were genotyped with two markers and also evaluated for CMD resistance in a high CMD pressure area in Nigeria. Results of the marker analysis and phenotypic evaluation of CMD resistance in the field revealed that the markers RME1 and NS158 SSR were excellent prediction tools for CMD resistance in some crosses (a prediction accuracy of 70–80 percent). In a few families, however, the markers were not polymorphic between the resistant and susceptible parent and, therefore, were not useful. This highlights the need to develop many markers around a gene of interest in a MAS programme and then to use those markers to evaluate the parents and identify the best markers for the different cross combinations.

Eighteen progenies from TME3 carrying the *CMD2* marker were established from embryo axes and imported to CIAT from IITA.[1] They were crossed extensively to elite parents. Seeds harvested from the crosses were germinated *in vitro* from embryo axes according to standard protocols for cassava (Fregene *et al.,* 1997, CIAT, 2002) to allow sharing the CMD resistant genotypes with collaborators in Africa and India. Each plantlet was multiplied after three to four weeks of growth to obtain three to five plants. After another four weeks, leaves of all plants were removed for molecular analysis and the plants multiplied again to obtain 10–20 plantlets. DNA isolation was by a rapid mini preparation method developed for rice (Nobuyuki *et al*., 2000). The DNA obtained is sufficient for 100 reactions and can be held in the Costar plates for two months at –20 °C without any degradation. PCR amplification, polyacrylamide gel electrophoresis (PAGE) or agarose gel analysis of SSR markers NS158 and RME1 were as described by Mba *et al*. (2001). The versatility of spreadsheets makes them the appropriate software to handle the diverse information generated by MAS. Gel images from the marker analysis were entered directly into a spreadsheet that contains information on the parents, tissue culture and greenhouse records, and subsequent phenotypic evaluation of the progenies. After molecular analysis, genotypes that carry the marker allele associated with *CMD2* were further multiplied to obtain at least 30 plants. Ten plants were sent to the greenhouse for hardening and later transferred to the breeding programme for evaluation. Five plants were kept *in vitro*, while 15 plants were shipped to partners in India and Africa as shown in the flow chart for MAS (Figure 4).

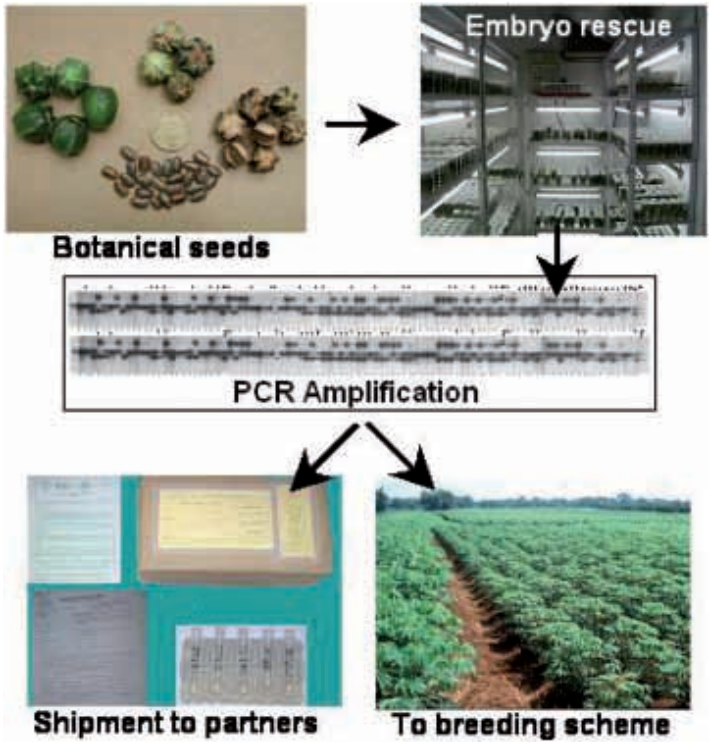
To date, more than 50 000 progeny have been evaluated with CMD linked markers and resistant lines shared with national programmes in India or Africa, and also incorporated into the breeding scheme at CIAT. The cost of a single marker data point is US$0.30 and 32 000 samples can be processed in a year.

MAS for CMD resistance at a NARS

Although evaluation for CMD resistance in sub-Saharan Africa is relatively easy and most areas have sufficient disease pressure to permit moderate to high heritability

[1] Phytosanitary conditions for the exchange of cassava germplasm between Africa and Asia are very stringent, but appropriately indexed *in vitro* cultures of embryo axes are permitted for experimental purposes.

FIGURE 4
Schematic representation of steps employed in breeding for resistance to the CMD in Latin America cassava gene pools

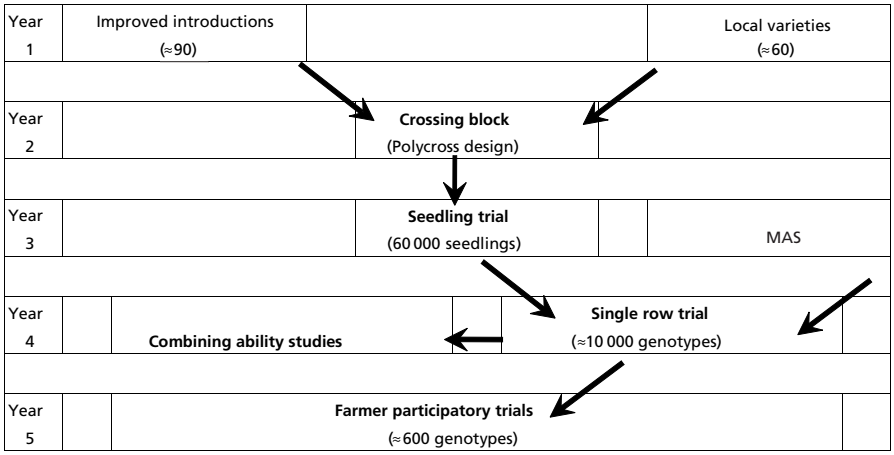


The entire process from sexual seeds to tissue plants for shipment or transfer to the screen house takes approximately three months.

of resistance, overlapping outbreaks of CGM, cassava bacterial blight (CBB), and CMD are common (Legg and Ogwal, 1998) and the need for modest-sized breeding populations make MAS for CMD resistance a powerful tool to accelerate cassava improvement even in Africa. A MAS and participatory plant breeding (PPB) project was initiated in 2003 with funding from the Rockefeller Foundation to improve the resistance of local cassava varieties in the United Republic of Tanzania to CMD and CGM and also to provide proof of concept for the use of MAS to accelerate cassava improvement. The United Republic of Tanzania is the fourth largest producer of cassava in Africa with average yields of about 8 tonnes/ha (FAO, 2001). This is below the continent's average of 10 tonnes/ha, and well below the average yield of 14 tonnes/ha of Africa's (and the world's) largest producer, Nigeria.

The low yield in the United Republic of Tanzania is caused by many factors, including the susceptibility of commonly grown varieties to major diseases and pests such as CMD and the cassava brown streak disease (CBSD). The project crosses farmer-preferred germplasm, by agro-ecology, to improved introductions that are resistant to CMD and to CGM. Markers associated with resistance to CMD are used to reduce

FIGURE 5
MAS scheme to improve local varieties of cassava in the United Republic of Tanzania using improved disease and pest resistant introductions from Latin America



The scheme is now in its second year.

the population size and a small set of genotypes with the "minimum criteria" for successful cassava production are evaluated in a single season in the corresponding agroecology and then evaluated over two cycles in collaboration with end-users (rural communities and cassava processors). Figure 5 describes the scheme of the United Republic of Tanzanian MAS and PPB project. CMD resistant F_1 generated by MAS at CIAT were crossed to BC_1 derivatives of *M. esculenta* sub spp. *flabellifolia*, showing good resistance to CGM, to produce progenies that combine some CMD and CGM resistance (Kullaya *et al.*, 2004). The progenies were established from embryo axes as *in vitro* plants to aid shipment to Africa. Molecular markers associated with resistance to CMD and phenotypic evaluation for CGM resistance were used to screen and select progenies that combine resistance to CMD and CGM. Resistant plants (300 genotypes and ten plants per genotype), were shipped to the United Republic of Tanzania as *in vitro* plantlets for use as improved parents. A selection based on harvest index, a highly heritable trait, and total biomass was made and 80 genotypes selected. These were planted in the second year in a controlled crossing block together with 54 local germplasm from the eastern and southern zones of the country. Emphasis was placed on local varieties with, or tolerance to, CBSD, which is a major disease of cassava in coastal east Africa and Mozambique. Over 40 000 crosses were made between the improved genotypes and the local varieties producing more than 60 000 seeds.

Sexual seeds obtained from crossing improved and local genotypes were planted in the screen house and transferred to the field 40 days after planting. Parental lines were also planted in the screen house from woody stakes. DNA was isolated from parental lines using the rapid mini-preparation method and evaluated with the five markers associated with the

CMD2 mediated resistance to CMD. Polymorphism in pair-wise combinations of the parental lines was observed with at least one of the five markers and will be used on the progeny. The phenotype of the progeny will be evaluated at three and six months after planting for resistance to CMD, CBSD and CGM. Markers are currently being tested for CGM resistance and are being developed for resistance to CBSD; when their utility is confirmed, they will also be used on progenies.

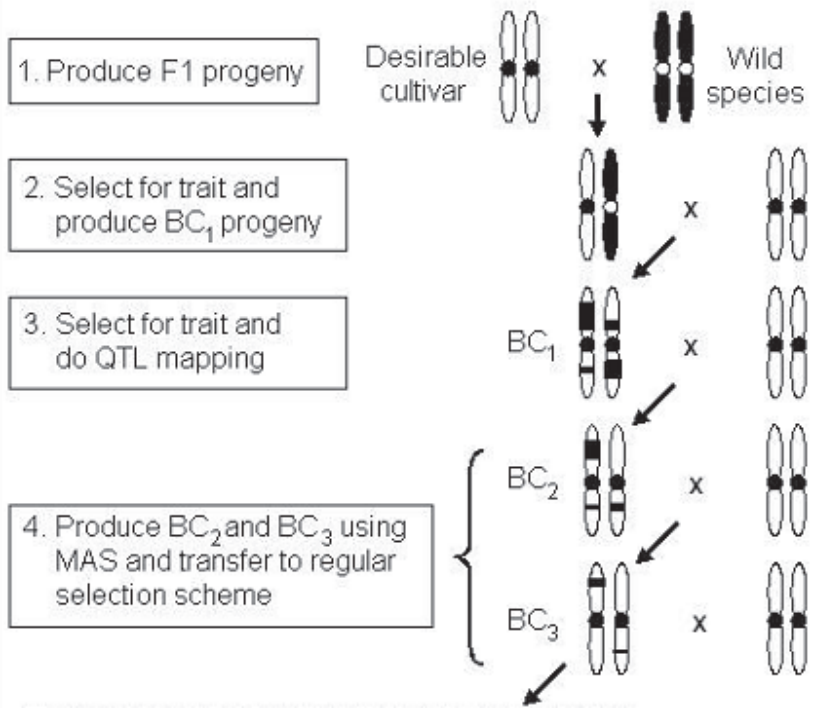
Using published broad sense heritability of 0.6 for CMD resistance (Hahn, Terry and Leuschner, 1980), it is expected that 24 000 symptomless genotypes will be analysed with markers associated with resistance to CMD. The gain of MAS will be the elimination of at least 38 400 (4 800 x 8 plants) that would have been carried to the single row trial stage (eight plant-rows per genotype), considering that breeders traditionally select 20 percent at the seedling trial stage. This represents a reduction of about 4 ha at the CET. If markers can be used to select for resistance to CGM and CBSD, then an additional number of genotypes can be eliminated from the CET leading to even greater savings. Using MAS for CMD alone would reduce the size of field trials by 50 percent. If additional second and third traits were included, reductions could be as high as 75 and 87.5 percent, respectively. Perhaps the most important advantage, however, comes from the increased genetic gain arising from higher heritabilities in these field evaluations with fewer genotypes.

MAS for transferring useful traits from wild relatives of cassava into the cultivated gene pool

Wild *Manihot* germplasm offers a wealth of useful genes for the cultivated *M. esculenta* species, but its use in regular breeding programmes is restricted by linkage drag and the long reproductive breeding cycle. For example, several accessions of *M.esculenta* sub spp. *flabellifolia*, *M. peruviana* and *M. tristis* have high levels of proteins (Nichols, 1947; Asiedu *et al*., 1992; CIAT, 2004). Low amylose content starch (3–5 percent) or waxy starch of relevance to the cassava starch industry has also been identified in two wild relatives of cassava, namely *M. crassisepala* and *M. chlorostricta.* The only source of dramatically delayed post-harvest physiological deterioration (PPD) has been identified in an interspecific hybrid between cassava and *M. walkerae*. The *M. walkerae* parent was collected in Mexico and held at the Washington University, St. Louis, United States of America (Bertram, 1993). It was brought to CIAT in 1998 in an attempt to use it in improving PPD. Furthermore, the only source of resistance to the cassava hornworm and the most widely deployed source of resistance to CMD were identified in fourth backcross generation progenies of *M. glaziovii* (Jennings, 1976; Chavarriaga *et al.,* 2004). Moderate to high levels of resistance to CGM, whiteflies and the cassava mealybug have been found in interspecific hybrids of *M. esculenta* sub spp. *flabellifolia.* The delayed PPD trait and resistance to the pests were successfully transferred to F_1 interspecific hybrids suggesting dominant or additive gene action of the gene(s) involved (CIAT, unpublished data).

The long reproductive cycle and lengthy time required to develop new cassava varieties (10–15 years) often discourages the use of wild species in most conventional cassava breeding programmes. However, the use of molecular markers to introgress a single target region of the genome can

FIGURE 6
Modified advanced backcross QTL scheme in the introgression of useful traits from wild relatives into cassava



save between two to four backcross generations (Frisch *et al.*, 1999). Indeed, it has been shown in several crops that the "tremendous genetic potential" locked up in wild relatives can be released more efficiently through the aid of new tools of molecular genetic maps and the advanced backcross QTL mapping scheme (ABC-QTL) (Tanksley and McCouch, 1997).

For several years now molecular marker tools and a modified ABC-QTL scheme have been tested in cassava at CIAT for the introgression of useful genes from wild relatives. The scheme entails generating BC_1 crosses and carrying out QTL mapping followed by selection of genotypes carrying the genome region of interest with minimum segments of the donor genome (Figure 6). The modified ABC-QTL is currently being used at CIAT to introgress genes for high protein content, waxy starch, delayed PPD, and resistance to whiteflies and the hornworm. The most advanced of these MAS projects is the introgression of high protein content from close wild relatives of cassava. Two BC_1 families of between 250 and 300 progenies were developed from two accessions of *M. esculenta* sub spp. *flabellifolia* OW284-1 and OW231-3, and the improved cassava variety from Thailand Rayong 60 (MTAI 8 in the germplasm collection). The BC_1 families were planted

in a CET for evaluation of root protein content at ten months. The grand parental lines of the BC_1 population were genotyped with over 800 simple sequence repeat (SSR) markers available for cassava and about 300 polymorphic markers were identified. The polymorphic markers are being assayed in the progenies after which QTL analysis will be conducted using the phenotypic protein and molecular marker data. Genotypes that have QTL for protein and a minimum of the donor parent genome will be selected and used for producing the BC_2 generation.

For introgression of naturally occurring mutant granule-bound starch synthase (GBSSI) for waxy starch in wild relatives, a more targeted approach was taken. Sequencing of the glycosyl transferase region of the GBSSI gene from the wild relatives and two cassava accessions identified four single nucleotide polymorphisms (SNPs) that differentiated the wild accessions from cassava. Allele-specific molecular markers unique to these SNPs were developed for selection of these alleles in a breeding scheme.

Genetic crosses were made between *M. chlorosticta* accession CW14-11 and MTAI8, and the resulting F_1 was back-crossed to MTAI8. The allele specific marker will be used together with other agronomic traits, particularly performance, to select for BC_1 that carry the mutant GBSS alleles for self-pollination to recover the waxy trait. The identification of natural mutants in a key gene and development of markers represent an innovative molecular tool to accelerate the introgression of favourable alleles from wild relatives into cassava. Backcross derivatives have also been developed from *M. walkerae* (MWal 001) for delayed post-harvest physiological deterioration; from MNG11 (a BC_4 derivative of *M. glaziovii)* for resistance to hornworm; and from *M. esculenta* sub spp. *flabellifolia* (FLA447-1) for resistance to whiteflies. Phenotypic and genetic mapping of these backcross populations are in progress to be followed by identification of QTL and selection of progenies to generate the next generation. MAS will later be used to combine these genes into progenitors for use as parents in breeding which, together with low cost marker technologies, will be distributed extensively to national programmes in Africa, Asia and Latin America to produce improved varieties.

Marker-assisted estimation of average heterozygosity during inbreeding of cassava

A principal use of molecular markers by private sector breeding companies is to accelerate the development of inbred lines. Cassava genotypes are heterozygous and very little inbreeding has been practised to date. However, inbred lines are better as parents as they do not have the confounding effect of dominance and carry lower levels of genetic load (undesirable alleles). Speed of inbreeding depends upon the average heterozygosity of the original parental lines, the homozygosity level of the selected genotypes at the end of the self-pollinating phase and the process of selection of progenies to be self-pollinated (Scotti *et al.,* 2000). Basically in the inbreeding process two events go together: phenotypically there is a decrease in vigour, which is correlated with the increased levels of homozygosity. While the aim is to select vigorous plants (tolerant to inbreeding), in the process plants may be selected that are less homozgygous than the expected average for their generation. It is expected that the first few cycles of self-pollination will result in a marked reduction of vigour (inbreeding depression associated with the genetic load of the parental lines); therefore, selection for

tolerance to inbreeding depression must be exerted. However, such selection is biased by the differences in homozygosity levels of segregating partially inbred genotypes. This highlights the need for a method to measure the level of heterozygosity in these partially inbred individuals and to use this in a co-variance correction in the selection of phenotypically vigorous genotypes. Molecular markers can be used to estimate the level of homozygosity of a given plant, enabling selection of plants with true tolerance to inbreeding.

Molecular markers can identify regions in the genome that are particularly related to the expression of heterosis and for measuring genetic distances among inbred lines to direct crosses with higher probabilities of high heterosis. Co-dominant SSR markers on a genome-wide basis are suitable for this purpose. The effect of self-pollination on vigour and heterozygosity was analysed in nine S_1 families, heterozygosity being estimated in the S_1 families by 100 mapped SSR markers that cover over 80 percent of the cassava genome and plant vigour by dry root yield and plant biomass. Results will assist in selecting the best performing and least heterozygous plants during inbreeding by identifying superior partially inbred parental lines. Molecular markers could also be used to delineate heterotic groups in cassava. Genetic resources of cassava have been characterized at the regional (Fregene *et al.*, 2003) and global (Hurtado *et al.*, 2005) levels. Highly differentiated groups of accessions were observed particularly among groups of materials from Guatemala and Africa and they may represent heterotic pools. These groupings are being tested based on molecular markers by genetic crossing between and within the groups as a first step to define heterotic patterns for a more systematic improvement of combining ability via recurrent reciprocal selection.

Other potential MAS targets

Several other traits for which MAS can be applied to increase efficiency of breeding include:

Beta-carotene

CIAT and a number of partners are involved in a project to produce cassava varieties with higher levels of β-carotene in yellow roots. This is one way of combating the deficiency of this key micronutrient in areas where cassava is a major staple. The experimental approach to increasing cassava β-carotene content includes conventional breeding and genetic transformation. The discovery of a wide segregation pattern of root colour in two S_1 families from the Colombian landrace MCOL 72 (cross code AM 273) and MTAI 8 (AM 320) was the basis for molecular genetic analysis of β-carotene content in cassava. Three markers, SSRY251, NS980 and SSRY330, were found to be associated with β–carotene content. These are in the same region of the genome and together explain >80 percent of phenotypic variation for β-carotene content in the population used for this study. The homozygous state of certain alleles of these markers translates into higher β-carotene content, suggesting that breeding for this trait can benefit from molecular markers to assist in combining favourable alleles in breeding populations. The work is continuing with the search for additional favourable alleles in yellow-rooted germplasm to give the best possible phenotypic expression of the trait.

Cyanogenic potential

A collaborative project between the Swedish University of Agricultural Sciences (SLU), Uppsala, the Medical Biotechnology

Laboratories (MBL), Kampala, and CIAT, is aimed at the genetic mapping of CNP in cassava. An S_1 family-AM 320, derived from the bitter variety MTAI 8 is the basis for the study. This family has been evaluated for cyanogenic glucoside content and has been genotyped with more than 200 diversity array technology (DarT) markers at CAMBIA, Australia, and 150 SSR markers at CIAT. The discovery of molecular markers for CNP will provide a tool to select efficiently for low cyanogenic potential in cassava. Also ongoing is the genetic mapping of the two cytochrome P450 genes CYP79D1 and D2 that catalyse the rate-limiting step of the biosynthesis of the cyanogenic glucosides, linamarin in the S_1 family AM 320. The group is also looking for an association with QTL for CNP. It is expected that markers associated with CNP will be identified at the end of the study.

Dry matter content

Few key traits in cassava hold greater potential for increasing cost-effectiveness via MAS than root dry matter content (DMC). This trait is usually measured at the end of the growth cycle. A number of genetic and environmental effects influence DMC. It is usually highest before the onset of rains, but drops after the rains begin as the plant mobilizes starch from the roots for re-growth of leaves (Byrne, 1984). Defoliation from pest and disease attacks can lower DMC. Breeding programmes have been quite successful in improving DMC, especially for industrial markets. The entry point for developing markers associated with DMC was recent diallel experiments (Jaramillo *et al.*, 2005; Calle *et al.*, 2005; Pérez *et a*l., 2005a, b; Cach *et al*., 2005b). Diallels, in this case made up of 90 families, are an ideal method to identify genes controlling DMC that are useful in many genetic backgrounds. Estimates of general and specific combining ability (SCA and GCA, respectively) for many traits of agronomic interest were calculated, with emphasis on DMC. Based on GCA estimates, parents were selected to generate larger-sized progenies for DMC mapping. Sizes of families in the original diallel experiment were about 30 progenies, which is rather small for genetic mapping. Parallel to the development of mapping populations was the search for markers associated with DMC using two F_1 families, GM 312 and GM 313, selected from the diallel experiment having parents with high GCA for DMC.

Initial marker analysis using bulked segregant analysis led to the discovery of two molecular genetic markers, SSRY160 and SSRY150, which explain about 30 and 18 percent, respectively, of phenotypic variance for DMC. These markers are being analysed on approximately 700 genotypes derived from 23 crosses with parents having high GCA for DMC in order to confirm their utility across genetic backgrounds. Parallel to this, larger families are being developed from selected parents for QTL mapping of DMC.

Disadvantages of MAS

Perhaps the greatest disadvantage of MAS is the time and financial investment required to develop markers that are widely applicable for traits of agronomic importance. Often a marker developed in one or a few related genotypes will not work for other genotypes in a breeding scheme due to allelic effects. Furthermore, development of markers, particularly for QTL, is complicated by epistatic interactions and the critical need for good quality phenotypic data. Several ways around this

problem have been proposed, such as the use of candidate genes involved in the traits directly as selectable markers without the need for laborious gene tagging experiments. However, unravelling the genetics and the development of markers for such traits is still many years down the road. New methods of association mapping and linkage disequilibrium mapping that rely upon non-random association of candidate genes or markers on a high resolution map with a phenotype of interest in a non-structured collection of genotypes have been used extensively in human medicine to identify genes involved in disease (Cardon and Bell, 2001). Given the enormous difficulties of quantitative mapping in humans and the success of association mapping, these methods have also been proposed as ways around the problems in developing markers for low heritability traits in plants (Gaut and Long, 2003). The development of (partially) inbred cassava genetic stocks will certainly accelerate the application of MAS for the genetic improvement of the crop.

CONCLUSIONS

Given limited resources, further prioritization of traits is needed for the development of markers if they do not already exist. Top priority should be given to MAS for the most important pests and diseases prevalent in the region for which durable sources of resistance genes exist. Priority should also be given to DMC as this is another trait that, although having a high narrow sense heritability at the time of evaluation (usually after the onset of the rains to permit planting immediately thereafter), is significantly affected by non-genetic factors and is not as highly heritable. There are several initiatives to assist national programmes acquire new molecular tools to increase the cost-effectiveness of breeding. Prominent among these are the "molecular breeding communities of practice" project of the Generation Challenge Programme (GCP, www.generationcp.org) and the Rockefeller Foundation-funded African Molecular Marker Network (AMMANET, www.africancrops.net/ammanet). Both have training programmes on molecular breeding that are open to national programme scientists. The CIAT cassava project has also developed a Web-based database resource including protocols, populations, and markers for MAS in cassava that can easily be accessed by national programmes (www.ciat.cgiar.org/mascas).

Cassava and common beans: contrasts

Cassava and beans are similar with respect to the modest level of research input they have enjoyed over the past three to five decades. Both have been part of the research agenda of CIAT and of the CGIAR for nearly thirty years, and especially beans have benefited from inputs from laboratories and programmes in the United States of America and, to a lesser degree, Europe. However, research investments for high-scale genomics through marker development in these crops has been far less than for the "super crops" like maize, rice or soybean that enjoy participation by the private sector, but are more than minor orphan crops with local usage in the tropics.

Yet biologically, these two crops are widely contrasting. Cassava is a perennial versus beans, which are short-season annuals, although climbing beans at high altitudes can be similar to cassava in growth cycle. Beans are an autogamous seed crop while cassava is an allogamous crop with vegetative propagation. Accompanying this latter dichotomy are differences in gene action. Beans present largely additive gene

action, while cassava expresses important components of dominance and epistatic action. Finally, cassava as a clonal crop can fix heterotic combinations, while a lack of genetic male sterility or apomixis systems in common bean have curtailed the development of a hybrid industry for this seed crop even though heterosis is observed.

In spite of their biological and other differences, the results of several years experience with MAS in beans and cassava are surprisingly similar. In both crops, MAS is being employed principally to bolster phenotypic selection for disease resistance genes. Disease resistance is often governed by relatively few genes, and phenotypic data are obtained more easily. On the other hand, MAS for more complex traits has yet to find ready application. While there are candidates for such traits in both crops (root bulking in cassava; low phosphorus or drought tolerance in beans), the complexity of these traits has made the identification of reliable markers more difficult and has delayed application. Obtaining reliable phenotypic data for complex traits is especially difficult and is often the biggest bottleneck to eventual application of MAS. In the case of cassava, no inbred parents have been used to date for the development of molecular markers, making the genetic analysis more difficult.

However, some differences in the application of MAS for the two crops may be noted, arising from the form of reproduction of each crop. The time frame to select cassava clones through multilocational trials is about six to seven years. During this period and with each step the number of genotypes is reduced as a result of the selection exerted, but the genotype of each individual clone remains stable. In the case of beans from the F_1 until stabilization of pure lines there is an intense segregation process in the early generations which tapers off in later generations. In both crops MAS can be used in the early stages of the selection process but with different objectives. In cassava, MAS can help to select early on the clone that will ultimately be released, whereas in beans MAS is used to "direct" the segregation process in the more desirable direction. Although maps with significant saturation are available for both crops, these have been constructed over several years, employing genotypes (in the case of beans) from different gene pools with wide polymorphism. A small proportion of these markers (often 20–30 percent) is polymorphic in other hybrid combinations among the genotypes within the same gene pool or race that have been created to tag a specific trait. Thus, genome coverage is often still not optimal for the high quality QTL analysis that is usually needed for complex traits.

RECOMMENDATIONS

Careful prioritization of traits, marker system and genetic stocks for MAS

The limited resources available for cassava or bean research require a judicious allocation of efforts. In the past 10–20 years there has been increased investment in molecular marker research in both crops. However, a considerable proportion of that research was directed at demonstrating the usefulness of different techniques, e.g. RAPD, restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), etc. Over this period there has been an ever-changing set of technologies but relatively little actual benefit derived from their application. There is a trade-off between being on the cutting edge with the newest technologies and "sticking it out" with an "outdated" technology

until some benefit is extracted from it. On the other hand, far too much effort has been expended in the identification of markers for traits without carrying these through to application. Often gene tagging is a component of a short-term project, and does not receive the necessary follow up in implementation. In each case, the essential question is: what are the key genes for each crop? And once defined, which genes merit the investment to develop molecular markers? For investments in molecular marker development to yield results, it is important that traits are chosen for which molecular breeding has both a clear advantage over field-based selection and is feasible in the short to medium term. It is also important that emphasis be given to selecting the relevant crosses, pedigrees and populations in which to practise MAS, and to have in place appropriate phenotyping strategies for the confirmation of MAS results. In this regard, the use of parental surveys of many of the genotypes involved in a given breeding programme is an important first step in implementing MAS.

Short- and long-term research related to MAS

The present research structure that is normally based on short-term projects, usually of three years' duration, can seldom be expected to deliver results of usable markers for complex traits. Such short-term projects that seek to establish the basis for MAS or to implement selection should limit their objectives to simply inherited traits. On the other hand, longer-term funding either of a programmatic or successive project funding nature, must be obtained to address more complex traits governed by QTL as these would normally require at least two phases of three-year projects. The first phase might be expected to reveal the inheritance of a given trait, establishing the location and numbers of QTL, while a second phase would be required to validate these over more environments and to find markers that are polymorphic over a wide number of genotypes and therefore widely useful for breeding, as well as adapted to rapid laboratory techniques. A medium- to long-term investment likewise implies careful prioritization of such traits, with regard to potential impact and the eventual need for MAS. These reflections are based upon presently available laboratory techniques, but as techniques for more detailed and widespread evaluation of loci and genotypes are developed (e.g. gene chips for analysis of multiple loci), conclusions could change significantly.

Scaling-up technologies

After the development of molecular markers for a trait and their initial implementation, a period of scaling-up in use of the specific markers is necessary. Sometimes this involves changes to MAS protocols, in the marker detection technique or in the markers themselves. Marker re-design has been a common element of scaling-up exercises and can involve something as simple as changing a PCR fragment size to implementing a SNP assay for the actual sequence differences between alleles. Technologies that speed up the implementation process and lower the costs associated with scaling-up are crucial to the success of MAS and are often neglected.

Development of markers that are useful in a large number of crosses

Often a marker developed for a particular trait in one or a few related genotypes will not work for other genotypes with high value of the trait due to differences in gene or allelic effects. Unravelling the genetics of

major traits of agronomic interest even in a subset of elite parents used for breeding is beyond the resources available for bean and cassava research. Association mapping and linkage disequilibrium mapping, which rely upon non-random association of candidate genes or markers on a high resolution map with a phenotype of interest in a non-structured collection of genotypes, have been proposed as a way around this problem. Association mapping can be used to discover new marker-trait associations or to validate associations that were found through conventional genetic mapping. The GCP is facilitating association mapping of traits of agronomic importance in cassava and beans with the goal of discovering more useful markers for a wider range of genotypes.

The need to strike a balance between MAS and field-based selection

Occasionally the question is raised: which is better, MAS or conventional selection? This very question betrays a false dichotomy that hinders progress. By itself, MAS is seldom an adequate selection tool and therefore must be combined with conventional phenotypic selection. The objective should be to develop the optimal balance between conventional and molecular breeding, and the "best" balance will be unique to each situation, crop, selection scheme, environment and opportunities for different selection methods. More emphasis is needed on combined selection systems, rather than viewing MAS as a replacement for phenotypic or field selection.

REFERENCES

Afanador, L.K. & Hadley, S.D. 1993. Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean. *Bean Improv. Coop*. 35: 10–11.

Akano, A.O., Dixon, A.G.O., Mba, C., Barrera, E. & Fregene, M. 2002. Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor. Appl. Genet.* 105: 521–525.

Asiedu, R., Hahn, S.K., Bai, K.V. & Dixon, A.G.O. 1992. Introgression of genes from wild relatives into cassava. *In* M.O. Akoroda & O.B. Arene, eds. *Proc. 4th. Triennial Symp. Internat. Soc. Trop. Root Crops-Africa Branch,* pp. 89-91. Nigeria, ISTRC-AB/IDRC/CTA/IITA.

Awale, H.E. &. Kelly, J.D. 2001. Development of SCAR markers linked to Co-42 gene in common bean. *Ann. Rep. Bean Improv. Coop*. 44: 119–120.

Babu, L. & Chatterjee, S.R. 1999. Protein content and amino acid composition of cassava tubers and leaves. *J. Root Crops* .25: 163–168.

Beaver, J.S., Rosas, J.C., Myers, J., Acosta, J., Kelly, J.D., Nchimbi-Msolla, S.M., Misangu, R., Bokosi, J., Temple, S., Aranud-Santana, E. & Coyne, D.P. 2003. Contributions of the bean/cowpea CRSP to cultivar and germplasm development in common bean. *Field Crops Res*. 82: 87–102.

Beebe, S., Velasco, A. & Pedraza, F. 1999. Marcaje de genes para rendimiento en condiciones de alto y bajo fósforo en las accesiones de frijol G21212 y BAT 881. *VI Reunião Nacional de Pesquisa de Feijão*. Salvador, Brazil.

Beebe, S., Skroch, P.W., Tohme, J., Duque, M.C., Pedraza, F. & Nienhuis, J. 2000. Structure of genetic diversity among common bean landraces of Mesoamerican origin based on correspondence analysis of RAPD. *Crop Sci.* 40: 264–273.

Beebe, S., Rengifo, J., Gaitan, E., Duque, M.C. & Tohme, J. 2001. Diversity and origin of Andean landraces of common bean. *Crop Sci.* 41: 854–862.

Beebe, S.E., Rojas, M., Yan, X., Blair, M.W., Pedraza, F., Muñoz, F., Tohme, J. & Lynch, J.P. 2006. Quantitative trait loci for root architecture traits correlated with phosphorus acquisition in common bean. *Crop Sci.* 46: 413–423.

Bellotti, A.C. 2002. Arthropod pests. *In* R.J. Hillocks, J.M. Thresh & A.C. Bellotti, eds. *Cassava: biology, production and utilization,* pp. 209-235. Wallingford, UK, CABI Publishing.

Bellotti, A.C., Arias B., Vargas, V.O., Reyes H.J.A. & Guerrero, J.M. 2002. Insectos y ácaros dañinos a la yuca y su control. *In* B. Ospina & H. Ceballos, eds. *La yuca en el tercer milenio. Sistemas modernos de producción, procesamiento, utilización y comercialización,* CIAT Publication No. 327, pp. 160-203. Cali, Colombia.

Bertram, R.B. 1993. *Application of molecular techniques resources of cassava (*Manihot esculenta *Crantz, Euphorfiabeae): interspecific evolutionary relationships and intraspecific characterization.* University of Maryland, USA. (Ph.D. thesis).

Blair, M.W. & Beaver, J.S. 1993. Inheritance of bean golden mosaic resistance from bean genotype A 429. *Ann. Rep. Bean Improv. Coop.* 36: 143–144.

Blair, M.W., Iriarte, G. & Beebe, S.E. 2003. QTL analysis of an Andean advanced backcross population for yield traits derived from wild *P. vulgaris*. *Ann. Rep. Bean Improv. Coop.* 46: 7–9.

Blair, M.W., Pedraza, F., Buendía, H.F., Gaitán-Solís, E., Beebe, S.E., Gepts, P. & Tohme, J. 2003. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.) *Theor. Appl. Genet*. 107: 1362–1374.

Blair, M.W., Beaver, J.S., Nin, J.C., Prophete, E. & Singh, S.P. 2006. Registration of PR9745–232 and RMC-3 red-mottled dry bean germplasm lines with resistance to bean golden yellow mosaic virus. *Crop Sci.* 46: 1000–1002.

Broughton, W.J., Hernandez, G., Blair, M.W., Beebe, S.E., Gepts, P. & Vanderleyden, J. 2003. Beans (*Phaseolus* spp.)–model food legumes. *Plant & Soil* 252: 55–128.

Buendia, H.F., Beebe, S.E., Blair, M.W., Teran, H. & Pedraza, F. 2003. Identificación de marcadores moleculares asociados a genes de rendimiento en una población RC2F3.7 de frijol común *Phaseolus vulgaris* L. (DOR390 x G19892). *Fitotecnia Colombiana* 3: 57–64.

Buitrago, A.J.A. 1990. *La yuca en la alimentación animal*, p. 446. Cali, Colombia, CIAT.

Byrne, D. 1984. Breeding cassava*. Plant Breeding Rev.* 2: 73–134.

Cach, N.T., Lenis, J.I., Perez, J.C., Morante, N., Calle, F. & Ceballos, H. 2005a. Inheritance of relevant traits in cassava (*Manihot esculenta* Crantz) for sub-humid conditions. *Plant Breed.* 125(2): 177–182.

Cach, N.T., Perez, J.C., Lenis, J.I., Calle, F., Morante, N. & Ceballos, H. 2005b. Epistasis in the expression of relevant traits in cassava (*Manihot esculenta* Crantz) for sub-humid conditions. *J. Heredity* 96: 586–592.

Calle, F., Perez, J.C. Gaitán, W., Morante, N., Ceballos, H., Llano, G. & Alvarez, E. 2005. Diallel inheritance of relevant traits in cassava (*Manihot esculenta* Crantz) adapted to acid-soil savannas. *Euphytica* 144: 177–186.

Cardon, L.R. & Bell, J.I. 2001. Association study design for complex disease. *Nature Revs. Genet*. 2: 91–99.

Carvalho, G.A., Paular, T.J., Alzate-Marin, A.L., Nietsche, S., Barros, E.G. & Moreira. M.A. 1998. Herança da resistência da linhagem AND 277 de feijoeiro-comun à raça 63-23 de *Phaeoisariopsis griseola* e identificação de marcador RAPD ligado ao gene de resistência. *Fitopatol. Bras*. 23: 482–485.

Ceballos, H., Iglesias, C.A., Pérez, J.C. & Dixon, A.G.O. 2004. Cassava breeding: opportunities and challenges. *Plant Mol. Biol*. 56: 503–515.

Ceballos, H., Fregene, M., Lentini, Z.J., Sanchez, T., Puentes, Y.I., Perez, J.C., Rosero, A. & Tofiño, A.P. 2006. Development and identification of high-value cassava clones. *Acta Hort.* 703:63–70.

Chavarriaga, P., Prieto, S., Herrera, C.J., López, D., Bellotti, A. & Tohme, J. 2004. Screening transgenics unveils apparent resistance to hornworm (*E. ello*) in the non-transgenic, African cassava clone 60444. *In A.* Alves & J. Tohme, eds. Adding value to a small-farmer crop. *Proc. 6th. Internat. Scientific Mtg. Cassava Biotech. Network,* Book of Abstracts, p.4. Cali, Colombia, CIAT.

Chávez, A.L., Sánchez, T., Jaramillo, G., Bedoya, J.M., Echeverry, J., Bolaños, E.A., Ceballos, H. & Iglesias, C.A. 2005. Variation of quality traits in cassava roots evaluated in landraces and improved clones. *Euphytica* 143: 125–133.

CIAT (International Center for Tropical Agriculture). 1997. Assessing and enhancing agrobiodiversity through biotechnology. *Annual Report*, 1994-1997, pp. 63–65. Cali, Colombia.

CIAT. 2001. Bean improvement for sustainable productivity, input use efficiency and poverty alleviation. *Annual Report*, pp. 37–38. Cali, Colombia.

CIAT. 2002. Bean improvement for sustainable productivity, input use efficiency and poverty alleviation. *Annual Report*, pp. 104–107. Cali, Colombia.

CIAT. 2003. Bean improvement for sustainable productivity, input use efficiency and poverty alleviation.. *Annual Report*, pp. 32–33 and 67–75. Cali, Colombia.

CIAT. 2004. Bean improvement for sustainable productivity, input use efficiency and poverty alleviation, *Annual Report*, pp. 84 and 115–117. Cali, Colombia.

FAO. 2001. *Food Outlook*. December 1999. Rome.

Ferreira, C.F., Borém, A., Carvalho, G.A., Nietsche, S., Paula, T.J., de Barros, E.G. & Moreira, M.A. 2000. Inheritance of angular leaf spot resistance in common bean and identification of a RAPD marker linked to a resistance gene. *Crop Sci*. 40: 1130–1133.

Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavarriaga, P., Roca, W., Tohme, J., & Bonierbale, M. 1997. A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor. Appl. Genet.* 95: 431–441.

Fregene, M., Suarez, M., Mkumbira, J., Kulembeka, H., Ndedya, E., Kulaya, A., Mitchel, S., Gullberg, U., Rosling, H., Dixon, A. & Kresovich, S. 2003. Simple sequence repeat (SSR) diversity of cassava (*Manihot esculenta* Crantz) landraces: genetic structure in a predominantly asexually propagated *Crop Theor. Appl. Genet*. 107: 1083–1093.

Frei, A., Blair, M.W., Cardona, C., Beebe, S.E., Gu, H. & Dorn, S. 2005. Identification of quantitative trait loci for resistance to *Thrips palmi* Karny in common bean (*Phaseolus vulgaris* L.). *Crop Sci.* 45: 379–387.

Frisch, M., Bohn, M. & Melchinger, A.E. 1999. Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Sci*. 39: 1295–1301.

Gaut, B.S. & Long, A.D. 2003. The lowdown on linkage disequilibrum. *Plant Cell* 15: 1502–1506.

Gepts, P. 1988. A middle American and an Andean common bean gene pool. *In* P. Gepts, ed. *Genetic resources of* Phaseolus *beans,* pp. 375-390. Dordrecht, Netherlands, Kluwer Academic Publishers.

Gimelfarb, A. & Lande, R. 1994. Simulation of marker-assisted selection in hybrid populations. *Genetical Res*. 63: 39–47.

Hahn, S.K., Howland, A.K. & Terry, E.R. 1980. Correlated resistance to cassava to mosaic and bacterial blight diseases. *Euphytica* 29: 305–311.

Hahn, S.K., Terry, E.R. & Leuschner, K. 1980. Breeding cassava for resistance to cassava mosaic disease. *Euphytica* 29: 673–683.

Haley, S.D., Afanador, L. & Kelly, J.D. 1994. Identification and application of a random amplified polymorphic DNA marker for the I gene (potyvirus resistance) in common bean. *Phytopathology* 84: 157–160.

Hallauer, A.R. & Miranda Fo., J.B. 1988. *Quantitative genetics in maize breeding, second edition.* USA, Iowa State University Press.

Hershey, C.H. 1984. Breeding cassava for adaptation to stress conditions: development of a methodology, pp. 303-314. *Proc. 6th Internat. Symp. Soc. Trop. Root Crops,* Lima, Peru.

Hillocks, R.J. & Wydra, K. 2002. Bacterial, fungal and nematode diseases. *In* R.J. Hillocks, J.M. Thresh & A.C. Bellotti, eds. *Cassava: biology, production and utilization,* pp. 261–280. USA and UK, CABI Publishing.

Hospital, F. & Charcosset, M. 1997. Marker-assisted introgression of quantitative loci. *Genetics* 147: 1469–1485.

Hospital, F, Chevalet, C. & Mulsant, P. 1992. Using markers in gene introgression breeding programs. *Genetics* 132: 1119–1210.

Hurtado L., P.X., Ospina, C., Marín, J., Buitrago R.Ch., Fregene, M.A., De Vicente, M.C., Wongtiem, P., Killian, A. & Wenzl, P. 2005. Comparación de dos marcadores moleculares: Microarreglos-microsatélites para el análisis de la diversidad estructural de yuca Manihot esculenta Crantz [resumen] [CD-ROM]. In *Congreso Asociación Colombiana de Fitomejoramiento y Producción de Cultivos* (9, 2005, Palmira, Colombia). [Resúmenes] [CD-ROM]. Corporación Colombiana de Investigación Agropecuaria (CORPOICA), Cali, CO. p. 138.

Ikeda, N., Bautista, N.S., Yamada, T., Kamijima, O. & Ishii, T. 2001. Ultra-simple DNA extraction method for marker-assisted selection using microsatellite markers in rice. *Plant Mol. Bio. Reporter* 19: 27–32.

Islam, F.M.A., Basford, K.E., Redden, R.J., Gonzalez, A.V., Kroonenberg, P.M. & Beebe, S. 2002. Genetic variability in cultivated common bean beyond the two major gene pools. *Genetic Resources and Crop Evolution* 49(3): 271–283.

Jaramillo, G., Morante, N., Pérez, J.C., Calle, F., Ceballos, H., Arias, B. & Bellotti, A. C. 2005. Diallel analysis in cassava adapted to the midaltitude valleys environments. *Crop Sci.* 45: 1058–1063.

Jennings, D.L. 1976. Breeding for resistance to African cassava mosaic. African cassava mosaic. *Report Interdiscipl. Workshop* IDRC 071e, pp. 39–44. Muguga, Kenya.

Jennings, D.L. & Iglesias, C. 2002. Breeding for crop improvement, pp. 149–166. *In* R.J. Hillocks, J.M. Thresh & A.C. Bellotti, eds. *Cassava: biology, production and utilization*. UK, CABI Publishing.

Johnson, W.C., Guzman, P., Mandala, D., Mkandawire, A.B.C., Temple, S., Gilbertson, R.L. & Gepts, P. 1997. Molecular tagging of the *bc*-3 gene for introgression into Andean common bean. *Crop Sci.* 37: 248–254.

Jung, G., Skroch, P.W., Coyne, D.P., Nienhuis, J., Aranud-Santana E., Ariyarathne, H.M., Kaeppler, S.M. & Bassett, M. 1997. Molecular marker-based genetic analysis of tepary bean-derived common bacterial blight resistance in different development stages of common bean. *J. Am.. Soc. Hort. Sci*. 122: 329–337.

Kawano, K. 1980. Cassava. *In* W.R. Fehr & H.H. Hadley, eds. *Hybridization of crop plants,* pp 225-233. Madison, WI, USA, ASA and CSSA.

Kawano, K. 2003. Thirty years of cassava breeding for productivity- biological and social factors for success. *Crop Sci.* 43: 1325–1335.

Kawano K., Narintaraporn, K., Narintaraporn, P., Sarakarn, S., Limsila, A., Limsila, J., Suparhan, D., Sarawat, V. & Watananonta, W. 1998. Yield improvement in a multistage breeding program for cassava. *Crop Sci.* 38: 325–332.

Kelly, J.D. & Adams, M.W. 1987. Phenotypic recurrent selection in ideotype breeding of pinto beans. *Euphytica* 36: 69–80.

Kelly, J.D., Schneider, K.A. & Kolkman, J.M. 1999. Breeding to improve yield. *In* S.P. Singh, ed. *Common bean improvement in the twenty-first century. Developments in plant breeding*, Vol. 7, pp. 185–222. Dordrecht, Netherlands, Kluwer Academic Publishers.

Kelly, J.D, Gepts, P., Miklas, P.N. & Coyne, D.P. 2003. Tagging and mapping of genes and QTL and molecular-marker assisted selection for traits of economic importance in bean and cowpea. *Field Crops Res*. 82: 135–154.

Klimyuk, V.I., Carroll, B.J., Thomas, C.M. & Jones, J.D.G. 1993. Alkali treatment for rapid preparation of plant material for reliable PCR analysis. *Plant J.* 3: 493–494.

Kullaya, A., Mtunda, K., Kulembeka, H., Ferguson, M., Marin, J., Ospina, C., Barrera, E., Jarvis, A., Morante, M., Ceballos, H., Tohme, J. & Fregene, M. 2004. Molecular marker-assisted and farmer participatory improvement of cassava germplasm for farmer/market preferred traits in Tanzania. *6th. Internat. Scientific Mtg. Cassava Biotech. Network.* Cali, Colombia, CIAT.

Legg, J.P. & Ogwal, S. 1998. Changes in the incidence of African cassava mosaic virus disease and the abundance of its whitefly vector along south-north transects in Uganda. *J. Appl. Entom*. 122: 169–178.

Liao, H., Yan, X., Rubio, G., Beebe, S.E., Blair, M.W. & Lynch, J.P. 2004. Basal root gravitropism and phosphorus acquisition efficiency in common bean. *Funct. Plant Biol*. 31: 1–12.

Mahuku, G., Montoya, C., Henríquez, M.A., Jara, C., Terán, H. & Beebe, S. 2004. Inheritance and characterization of angular leaf spot resistance gene in common bean accession G 10474 and identification of an AFLP marker linked to the resistance gene. *Crop Sci.* 44: 1817–1824.

Mba, R.E.C., Stephenson, P., Edwards, K., Melzer, S., Nkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, J. & Fregene, M. 2001. Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards a SSR-based molecular genetic map of cassava**.** *Theor. Appl. Genet.* 102: 21–31.

Melotto, M., Afanador, L. & Kelly, J.D. 1996. Development of a SCAR marker linked to the *I* gene in common bean. *Genome* 39: 1216–1219.

Meuwissen, T.H.E., Hayes, B.J. & Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819–1829.

Miklas, P.N., Larsen, R.C., Riley, R. & Kelly, J.D. 2000a. Potential marker-assisted selection for *bc-1*^{2} resistance to bean common mosaic potyvirus in common bean. *Euphytica* 116(3): 211–219.

Miklas, P.N., Smith, J.R., Riley, R., Grafton, K.F., Singh, S.P., Jung, G. & Coyne, D.P. 2000b. Marker-assisted breeding for pyramided resistance to common bacterial blight in common bean. *Ann. Rep. Bean Improv. Coop*. 43: 39–40.

Miklas, P.N., Stone, V., Daly, M.J., Stavely, J.R., Steadman, J.R., Bassett, M.J., Delorme, R. & Beaver, J.S. 2000c. Bacterial, fungal, and viral disease resistance loci mapped in a recombinant inbred common bean population ('Dorado'/XAN 176). *J. Am. Soc. Hort. Sci.* 125: 476–481.

Miklas, P.N., Kelly, J.D., Beebe, S.E. & Blair, M.W. 2006. Common bean breeding for resistance against biotic and abiotic stresses: from classical to MAS breeding. *Euphytica* 147: 105–131.

Muñoz, L.C., Blair, M.W., Duque, M.C., Roca, W. & Tohme, J. 2004. Level of introgression in interspecific (*Phaseolus vulgaris* x *P. acutifolius*) congruity-backcross lines. *Crop Sci.* 44: 637–645.

Nichols, R.F.W. 1947. Breeding cassava for virus resistance. *East Afr. Agric. J.* 12: 184–194.

Nietsche, S., Borém, A., Carvalho, G.A., Rocha, C., Paula, T.J., de Barros, E.G. & Moreira, M.A. 2000. RAPD and SCAR markers linked to a gene coferring resistance to angular leaf spot in common bean. *J. Phytopathol.* 148: 117–121.

Nobuyuki, I., Bautista, N.S., Yamada, T., Kamijima, O. & Ishi, T. 2000. Ultra simple DNA extraction method for marker-assisted selection using microsatellite markers in rice. *Plant Mol. Biol. Reporter* 18: 1–6.

Okogbenin, E., Porto, M.C.M & Dixon, A.G.O. 1998. Influence of planting season on incidence and severity of African cassava mosaic disease in the subhumic zone of Nigeria. *In* I.J. Ekanayake, ed. *Root crops and poverty alleviation*, pp. 338–392. Ibadan, Nigeria, IITA.

Park, S.O., Coyne, D.P., Mutlu, N., Jung, G. & Steadman, L.R. 1999. Confirmation of molecular markers and flower color associated with QTL for resistance to common bacterial blight in common beans. *J. Amer. Soc. Hort. Sci.* 124: 519–526.

Pastor-Corrales M.A., Jara, C. & Singh, S. 1998. Pathogenic variation in, source of, and breeding for resistance to *Phaeoisariopsis griseola* causing angular leaf spot in common bean. *Euphytica* 103: 161–171.

Pérez, J.C., Ceballos, H., Jaramillo, G., Morante, N., Calle, F., Arias, B. & Bellotti, A.C. 2005a. Epistasis in cassava (*Manihot esculenta* Crantz) adapted to mid-altitude valley environment. *Crop Sci.* 45: 1491–1496.

Pérez, J.C., Ceballos, H., Calle, F., Morante, N., Gaitán, W., Llano, G. & Alvarez, E. 2005b. Within-family genetic variation and epistasis in cassava (*Manihot esculenta* Crantz) adapted to the acid-soils environment. *Euphytica* 145: 77–85.

Polston, J.E. & Anderson, P.K. 1997. The emergence of whitefly transmitted geminiviruses in tomato in the western hemisphere. *Plant Disease* 81: 1357–1369.

Rajendran, P.G., Ravindran, C.S., Nair, S.G. & Nayar, T.V.R. 2000. True cassava seeds (TCS) for rapid spread of the crop in non-traditional areas. Thiruvananthapuram, Kerala, India, Central Tuber Crops Research Institute. 22 pp.

Santana, G.E., Blair, M.W., Morales, F., Mahuku, G., Jara, C. & Castaño, M. 2004. Uso de tecnicas clasicas y avanzadas para identificar genotipos de frijol resistentes a antracnosis y mosaico común. *Fitotecnia Colombiana* 4: 44–54.

Sartorato, A., Nietsche, S., Barros, E.G. & Moreira, M.A. 1999. SCAR marker linked to angular leaf spot resistance gene in common bean. *Bean Improv. Coop.* 42: 23–24.

Schneider, K.A., Brothers, M.E. & Kelly, J.D. 1997. Marker-assisted selection to improve drought tolerance in common bean. *Crop Sci.* 37: 51–60.

Scotti, C., Pupilli, F., Salvi, S. & Arcioni, S. 2000. Variation in Vigot and in RFLP-estimated heterozygosity by selfing tetraploid alfalfa: new perspectives for the use of selfing in alfalfa breeding. *Theor. Appl. Genet.* 101: 120–125.

Shull, G.F. 1952. Beginnings of the heterosis concept. *In* J.W. Gowen, ed. *Heterosis,* pp. 14–48. Ames, IA, USA, Iowa State College Press.

Singh, S.P. 1999. Integrated genetic improvement. *In* S.P. Singh, ed. *Common bean improvement in the twenty-first century. Developments in plant breeding*, Vol. 7, pp. 133–166. Dordrecht, Netherlands, Kluwer Academic Publishers.

Singh, S.P., Gepts, P. &. Debouck, D.G. 1991. Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Econ. Bot*. 45: 379–396.

Singh, S.P., Cardona, C., Morales, F.J., Pastor-Corrales, M.A. & Voysest, O. 1998. Gamete selection for upright carioca bean with resistance to five diseases and a leafhopper. *Crop Sci.* 38: 666–672.

Stam, P., & Zeven, A.C. 1981. The theoretical proportion of donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. *Euphytica* 30: 227–238.

Storey, H.H. & Nichols, R.F.W. 1938. Studies on the mosaic disease of cassava. *Ann. Appl. Bio*. 25: 709–806.

Sullivan, J.G. & Bliss, F.A. 1983. Expression of enhanced seed protein content in inbred backcross lines of common bean. *J. Am. Soc. Hort. Sci.* 108: 787–791.

Swanson, M.M. & Harrison, B.D. 1994. Properties, relationships and distribution of cassava mosaic geminiviruses. *Trop. Sci.* 34: 15–25.

Tanksley, S.D. & McCouch, S.R. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* 277: 1063–1066.

Urrea, C.A., Miklas, P.N., Beaver, J.S. & Riley. R.H. 1996. A codominant randomly amplified polymorphic DNA (RAPD) marker useful for indirect selection of BGMV resistance in common bean. *J. Am. Soc. Hort. Sci*. 121: 1035–1039.

Vallejo, V. & Kelly, J.D. 2001. Development of a SCAR marker linked to Co-5 gene in common bean. *Ann. Rept. Bean Improv. Coop*. 44: 121–122.

Van Schoonhoven, A. & Pastor-Corrales, M.A. 1987. Standard system for the evaluation of bean germplasm. Cali, Colombia, CIAT.

Yan, X., Liao, H., Beebe, S.E., Blair, M.W. & Lynch, J.P. 2005. Molecular mapping of QTLs associated with root hairs and acid exudation as related to phosphorus uptake in common bean. *Plant & Soil* 265: 17–29.

Young, R.A., Melotto, M., Nodari, R.O. & Kelly, J.D. 1998. Marker assisted dissection of the oligogenic anthracnose resistance in the common bean cultivar G2333. *Theor. Appl. Genet*. 96: 87–94.

CHAPTER 8

Marker-assisted selection in maize: current status, potential, limitations and perspectives from the private and public sectors

Michel Ragot and Michael Lee

SUMMARY

More than twenty-five years after the advent of DNA markers, marker-assisted selection (MAS) has become a routine component of some private maize breeding programmes. Line conversion has been one of the most productive applications of MAS in maize breeding, reducing time to market and resulting in countless numbers of commercial products. Recently, applications of MAS for forward breeding have been shown to increase significantly the rate of genetic gain when compared with conventional breeding. Costs associated with MAS are still very high. Further improvements in marker technologies, data handling and analysis, phenotyping and nursery operations are needed to realize the full benefits of MAS for private maize breeding programmes and to allow the transfer of proven approaches and protocols to public breeding programmes in developing countries.

INTRODUCTION

The ability to identify genetic components of traits, particularly quantitative traits, in Mendelian factors, and to monitor or direct their changes during breeding through the use of DNA-based markers has created much enthusiasm. Claims were sometimes made that marker-assisted selection (MAS) would rapidly replace phenotypic selection and dramatically reduce the time required to develop commercial varieties (Mazur, 1995). At the turn of this century, phenotypic selection was still the approach on which maize breeding programmes mostly relied to develop new and improved cultivars while MAS had contributed to advances in introgression, or backcross breeding (Ragot *et al.,* 1995; Ho, McCouch and Smith, 2002; Ribaut, Jiang and Hoisington, 2002; Morris *et al.,* 2003). Overly optimistic statements and exaggerated promises about the power of MAS to improve complex traits created excessively high and largely unfulfilled hopes and prompted a wave of cautious and sometimes pessimistic views (Melchinger, Utz and Schön, 1998; Young, 1999; Goodman and Carson, 2000; Bernardo, 2001).

Recently, multinational corporations with large maize breeding programmes reported the routine and successful use of MAS (Johnson, 2004; Niebur *et al.,* 2004; Eathington, 2005; Crosbie *et al.,* 2006). Rates of genetic gain twice as high as those achieved through conventional breeding were reported for MAS in maize. Accounts were also given of a number of MAS-derived single-cross (i.e. simple) hybrids being currently on the market. Although too little is known about the methods (e.g. breeding schemes, mathematical algorithms) and tools (e.g. marker technologies, computer programs, databases) used to develop these hybrids, these results have raised confidence in the ability of MAS to increase the rate of genetic gain over what can be achieved through conventional breeding. As technologies evolve and marker genotypes become less expensive, MAS becomes increasingly within the reach of developing countries. Whenever necessary, transfer of methods or tools from private companies to developing countries should be made possible while preserving the commercial interests of the companies concerned, thereby contributing to increasing the rate of genetic gain where it is most needed.

Much has happened in maize breeding since Stuber and Moll (1972) first reported that selection for grain yield in maize had resulted in changes in allele frequencies at several isozyme loci throughout the genome. In so doing, they essentially laid the grounds for MAS in maize. Indeed, if phenotypic selection could produce a change in marker allele frequencies, then why could deliberately altering marker allele frequencies at specific loci not produce predictable phenotypic changes for one or several traits?

The objectives of this chapter are to provide the scientific community and decision-makers with information on the current status of MAS in maize breeding programmes, including the major steps that led to it, and to provide suggestions to developing countries for deploying the technology and methods involved in an efficient, cost-effective and realistic manner.

HOW HAS MAS BEEN USED BY THE PRIVATE SECTOR TO IMPROVE THE MAIZE CROP?

Applications of DNA markers in private maize breeding programmes started in the 1980s with the identification of DNA clones used to detect restriction fragment length polymorphisms (RFLPs)

in the nuclear genome. As described below, the methods used to detect RFLPs were incompatible with the magnitude, speed and efficiency of all but a few aspects of selection in maize breeding programmes. Gradually, however, the methods used to detect DNA polymorphisms and to create meaningful information from DNA marker and phenotypic data sets have evolved to the point where they are routine components of some maize breeding programmes in the private sector.

Selection occurs at various stages in maize breeding programmes. The first opportunity arises when choosing inbred lines to mate as parents of new populations. In some programmes, all such inbreds are genotyped systematically at DNA marker loci (Smith and Smith, 1992). If the marker loci are sufficiently close on genetic or physical maps then reasonably good inferences may be made about the inbred's haplotype. Such information is used to establish identity, resolve disagreements related to germplasm ownership and acquisition, enforce laws intended to encourage genetic diversity of the hybrids and avoid using inbreds that contain transgenes which may violate regulatory considerations and restrictions. These selection practices, while admittedly not conventional MAS, have led to improvements in the maize crop by enabling more informed stewardship and deployment of genetic resources and by providing a degree of protection of intellectual property and related investments in maize breeding.

Unquestionably, the most pervasive and direct use of MAS in maize by the private sector has been with backcrossing of transgenes into elite inbred lines, the direct parents of the commercial hybrids (Ragot *et al.,* 1995; Crosbie *et al.,* 2006). Currently, the most widely deployed transgenes and combinations thereof (i.e. gene stacks) are for resistance to herbicides or insects (e.g. *Ostrinia* and *Diabrotica*). As the commercial maize crop of any region, maturity zone, market or country is not yet uniform or homogeneous for any transgene, maize breeders have elected to develop near-isogenic versions (transgenic and non-transgenic) of elite inbreds and commercial hybrids in order to satisfy combinations of licensing agreements, agronomic practices, regulatory requirements, market demands and product development schemes. This has required companies to have two parallel maize breeding programmes, transgenic and non-transgenic. In this manner, marker-assisted backcrossing (MABC) of transgenes, and to a lesser degree, of native genes and quantitative trait loci (QTL) for other traits, has expedited the development of commercial hybrids.

More recently, marker-assisted recurrent selection (MARS) schemes and infrastructure have been developed for "forward breeding" of native genes and QTL for relatively complex traits such as disease resistance, abiotic stress tolerance and grain yield (Ribaut and Betrán, 1999; Ragot *et al.,* 2000; Ribaut, Jiang and Hoisington, 2000; Eathington, 2005; Crosbie *et al.,* 2006). Simulation studies suggested that MAS could be effective for such traits under certain conditions (Edwards and Page, 1994; Gimelfarb and Lande, 1994), but the initial empirical attempts at such selection were not successful (Stromberg, Dudley and Rufener, 1994; Openshaw and Frascaroli, 1997; Holland, 2004; Moreau, Charcosset and Gallais, 2004) except in the special case of sweetcorn (Edwards and Johnson, 1994; Yousef and Juvik, 2001). The success reported for sweetcorn is due to the fact that the genetic base of sweetcorn is extremely narrow relative to dent or flint maize; thus

predicted gains and extrapolations across populations are more reliable. Also, phenotypic analyses of many traits in a sweetcorn breeding programme are extremely expensive because they involve processing large volumes of grain; therefore, MAS would be relatively inexpensive and effective under such circumstances. However, subsequent developments in technology, refinements in analytical methods and improvements in experimental designs have been assembled into a process that has shown promise for some reference populations of dent maize (Ragot *et al.,* 2000; Johnson, 2004; Crosbie *et al.,* 2006) as improvement in grain yield from MAS often exceeded that from non-MAS approaches. Presumably, such results will lead to the development of new and superior inbred lines and commercial hybrids in a cost-effective manner. While the impact of such MAS has not yet been fully realized in the maize crop, the methods have been employed to various degrees by programmes in the private sector that have the necessary infrastructure.

The potential for MAS to contribute to improvements in the maize crop should increase in parallel with our understanding of the relationships among genomes, the environment and phenotypes. Candidate transgenes will be developed on a regular basis and their contributions to maize improvement will be realized in the most efficient manner with MAS. Likewise, the identification of candidate native genes and their gene products and functions, and of other DNA sequences (e.g. miRNA, matrix attachment and regulatory regions), will improve the power of methods such as association mapping and genome scans to assess their genotypic value in the context of defined reference populations of significance to maize breeding (Thornsberry *et al.,* 2001; Rafalski, 2002; Niebur *et al.,* 2004; Varshney, Graner and Sorrels, 2005). Beyond its use in MARS schemes, this information might make it reasonable to reconsider ideas such as methods for predicting hybrid performance that may have been limited by the amount and type of information and by the design of the experiment when they were initially evaluated (Bernardo, 1994).

METHODOLOGY AND DESIGN OF BREEDING PROGRAMMES SUPPORTED BY MAS

As expected, private sector maize programmes focus entirely on inbred-hybrid breeding schemes intended to develop elite inbred lines that enable the profitable production of commercial F_1 hybrids. To a large extent, MAS breeding programmes use the same designs and methods known to maize breeders for decades and generic descriptions of these have been published (Hallauer and Miranda, 1981; Sprague and Dudley, 1988; Bernardo, 2002). When MAS is included in the breeding programme, the significant differences are, of course, the availability of genotypic data at different stages of selection and some knowledge of the relationships between the genotypic and phenotypic data sets for the reference population(s) in the target environment(s).

In contrast to conventional breeding schemes, the methods and design of infrastructure needed to support MAS have been the areas of greatest change. In order to utilize MAS, companies had to make significant investments to assemble or modify various aspects of infrastructure such as methods to detect DNA polymorphism, manage information, or analyse and track samples, software to relate genotype with phenotype, and off-season or continuous nurseries. These components had to be integrated with each

other and with breeding activities, which meant that scientists needed to learn how and when MAS provided a comparative advantage over other methods.

MAS: enabling methods, tools and infrastructure

Perhaps the component of infrastructure in greatest need of development was related to the acquisition of genotypic data (i.e. DNA markers). Although the concept of associating markers with quantitative traits was not new (Sax, 1923), the discovery reported by Stuber and Moll (1972) was very significant. Stuber and Moll (1972) described for the first time associations between molecular markers and quantitative traits while previous associations had been based on morphological markers (Sax, 1923). The advantages of molecular over morphological markers soon became obvious and detailed descriptions of these advantages were published by Tanksley *et al.* (1989) and Stuber (1992).

Two of these advantages are of particular importance. First, molecular marker genotypes can usually be obtained from any plant tissue, even from young seedlings or kernels, while morphological markers frequently require the observation of whole, mature plants. Selection can therefore occur earlier in the plant's cycle when using molecular markers than when using morphological markers. The ability to conduct early selection, possibly before flowering, can have a tremendous impact on the rate of genetic gain of a breeding programme and therefore constitutes a very significant advantage of molecular over morphological markers.

Second, molecular markers are neutral markers. They are not affected by environmental or growing conditions. They are not affected by the genetic background either, nor do they affect phenotypes. The expression of morphological traits, by contrast, can be dependent on environmental or growing conditions. In addition, epistasic interactions are often observed among morphological marker loci or between morphological marker loci and the genetic background. These epistatic interactions prevent distinguishing all genotypes associated with morphological markers and further limit the number of morphological markers that can be studied simultaneously.

Although isozyme markers had many advantages over morphological markers, the lack of a sufficient number of polymorphic loci limited their use for MAS (Goodman *et al.,* 1980). Nevertheless, isozyme markers are still used for quality control during seed production.

RFLPs (Botstein *et al.,* 1980) are based on DNA polymorphisms detected through restriction nuclease digestions followed by DNA blot hybridizations. The abundance and high level of polymorphism of RFLPs, especially in maize, allowed the construction of extensive maize genetic maps (Helentjaris *et al.,* 1986; Burr *et al.,* 1988; Hoisington, 1989; Coe *et al.,* 1995; Davis *et al.,* 1999) as well as the identification and mapping of many QTL.

Being robust, reproducible and codominant, RFLPs are perfectly suited for genetic studies as well as for MAS applications. Their two main disadvantages are the large quantities of DNA required, and the difficulty to miniaturize and automate. Nevertheless, RFLPs were quickly adopted and represented the marker system of choice for many plant species including maize throughout the 1980s and during much of the 1990s.

The development of the polymerase chain reaction (PCR) (Saiki *et al.,* 1988) turned out to be a major breakthrough in

molecular marker technology. PCR-based markers require little DNA, allowing sampling of young seedlings and very early selection and thereby optimization of breeding schemes. PCR-based marker protocols are very amenable to automation and miniaturization and improvements to protocols resulted in considerable reductions in both cost and time required to produce data points. The first two PCR-based marker systems were random amplified polymorphic DNA (RAPDs), and amplified fragment length polymorphisms (AFLPs). Detailed descriptions and critical assessments of these two systems can be found in Welsh and McClelland (1990), Williams *et al.* (1990), Penner *et al.* (1993), Ragot and Hoisington (1993), Skroch and Nienhuis (1995) and Jones *et al.* (1997) for RAPDs, and in Vos *et al.* (1995), Jones *et al.* (1997) and Castiglioni *et al.* (1999) for AFLPs. They are also described in other chapters of this book.

Simple sequence repeats (SSRs) or microsatellites rapidly became the marker of choice in maize, almost entirely displacing RFLPs and previously developed PCR-based marker systems. Polymorphism of SSRs is due to variable numbers of short tandem repeats, often two or three base pairs in length and usually flanked by unique regions (Tautz, 1989). SSRs are very reproducible (Jones *et al.,* 1997) and co-dominant (Shattuck-Eidens *et al.,* 1990; Senior and Heun, 1993; Senior *et al.,* 1996) and are therefore very suitable for maize MAS applications.

Many additional variations of PCR-based marker systems have been developed and a thorough review can be found in Mohan *et al.* (1997).

All the DNA-based marker systems described to date are gel-based systems, a major constraint for automation. Single nucleotide polymorphisms (SNPs) (Lindblad-Toh *et al.,* 2000) can be revealed in many ways including allele-specific PCR, primer extension approaches, or DNA chips, all of which are not gel-based. SNPs can generally be scored as co-dominant markers, except in the case of insertion-deletion polymorphisms. Although allelic diversity at SNPs is usually limited to two alleles, this limitation can be offset by the abundance of SNPs and the analysis of haplotypes, combinations of genotypes at several neighbouring SNPs. Haplotype analyses increase informativeness (Ching *et al.,* 2002), although at some expense because two to four SNPs have to be genotyped where one SSR sufficed. SNP genotyping can be highly miniaturized and automated, thereby reducing the cost and allowing the production of very large numbers of data points. With genetic maps containing several thousand mapped SNPs, these have become the marker of choice for private maize MAS programmes.

DNA marker technology has been a dynamic and often expensive component of the infrastructure needed for MAS. For example, one corporation indicated having spent tens of millions of United States dollars to develop an automated system for detecting RAPDs, a technology that was never suited for MAS in a large maize breeding programme. Later, another corporation spent an even greater amount of money to acquire technology for matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) analysis of amplified DNA fragments. These technologies were either rapidly replaced or never used. Such decisions would have bankrupted most national maize programmes or a couple of centres belonging to the Consultative Group on International Agricultural Research (CGIAR). Fortunately, this area of infrastructure has matured somewhat

and become more stable so that start-up and operating costs, while still high for some programmes, are more predictable.

Statistical methods and related software have also been areas of significant development, especially for the detection and description of putative QTL. QTL, which are nothing more than associations between markers and traits, were first described using simple association tests between trait values and marker genotypes (Stuber and Moll, 1972). These tests consider each marker locus independently and neither require nor take advantage of the existence of genetic maps. Statistical methods have been developed that take advantage of the existence of genetic maps (see review by Manly and Olson, 1999). These statistical methods, simple interval mapping (Lander and Botstein, 1989) and composite interval mapping (Jansen, 1993; Zeng, 1993, 1994), test the existence of associations between hypothetical marker genotypes and trait values at several points in intervals between pairs of adjacent marker loci on the genetic map, allowing the positioning of QTL on these genetic maps. All of the previous methods are based on single QTL models. Other statistical methods have been developed that simultaneously test the presence of several QTL in the genome (Kao, Zeng and Teasdale, 1999).

Many software packages are available for QTL mapping and based on one or several of the statistical methods developed to date. No two packages are exactly alike and all have specific strengths and weaknesses with respect to particular situations, making it sometimes beneficial to use more than one package to perform QTL mapping analyses. The software packages most commonly used for QTL mapping in maize include QTL Cartographer (Basten, Weir and Zeng, 1994), MapQTL (van Ooijen and Maliepaard, 1996), and PLABQTL (Utz and Melchinger, 1996). All of these only handle bi-allelic populations, while MCQTL (Jourjon *et al.,* 2005) also performs QTL mapping in multi-allelic situations, including bi-parental populations made from segregating parents, or sets of bi-parental, bi-allelic populations.

More recently, methods based on Bayesian analysis (Jansen, Jannink and Beavis, 2003; Gelman *et al.,* 2004) and association (Varshney, Graner and Sorrels, 2005) or *in silico* mapping (Parisseaux and Bernardo, 2004) have been proposed as more powerful and refined approaches to assess the relationships between genotype and phenotype that are needed for MAS. Methods of Bayesian analysis should be less affected by the uncertainties of QTL effects and locations and produce better estimates of those parameters in MAS. Association mapping approaches are particularly useful to validate the relevance of genes and alleles in specific germplasm such as that used by maize breeders. *In silico* mapping takes advantage of the pedigree relationships among individuals to structure the population used to establish marker-trait associations. This approach, which is highly complex due to the population structure resulting from pedigree breeding, is particularly appropriate for maize where data across many years and environments are available for large sets of related individuals. Certainly, as the annotation of genomes gradually improves, such methods will be common components of breeding programmes. Currently, the applications of methods such as association mapping for MAS are hindered by the fact that a very low percentage of the genes in crop plants have a function assigned to them on the basis of direct experimentation. However, this impoverished situation

is being enriched through a variety of projects on functional genomics.

In sharp contrast to the many methods and software packages developed for QTL identification and mapping, little has been published for MAS. This paucity of information on MAS tools most likely reflects both the low level of activity in the public sector and the fully proprietary nature of developments in the private sector.

In parallel with advancements in DNA technology and statistical methods, private sector programmes have enhanced the capabilities and capacities of their continuous nurseries. Such nurseries have been used for decades by programmes in both the private and public sectors. In order to conduct MAS to its greatest advantage, continuous nurseries had to be managed, equipped and staffed in new ways so that the plants complete their life cycle as quickly as possible and that the genotypic data (and sometimes some phenotypic data) needed for MAS may be collected at each sexual generation. Three to four sexual generations per year may be completed at such nurseries.

These activities and the continuous collection of both genotypic and phenotypic data in the target environment and their integrated analyses create huge data sets that must be analysed quickly and related to other large extant data sets. Data management and bioinformatics for breeders have therefore become critical components of the infrastructure needed to use MAS. Prior to the advent of MAS, some large private breeding programmes had established a group of dedicated data managers to assist with research and marketing, and with the arrival of genomics and MAS the need for such dedicated specialists has increased greatly.

Once the basic infrastructure had been established to complement the activities of maize breeders, programmes were ready to implement several basic aspects of MAS; many of which are derived from well established methods and principles of maize breeding.

MAS-based breeding

Selection occurs at various stages in maize breeding programmes. The first opportunity for selection is the choice of inbred lines to mate as parents of new populations. Prior to the advent of DNA marker data, the selection of such parents would be based on a combination of phenotypic assessments, pedigree information, breeding records and chance (Hallauer and Miranda, 1981; Sprague and Dudley, 1988). In some programmes today, all such inbreds are genotyped systematically at DNA marker loci. Depending on the resources and objectives, the degree of genotyping may range from a low density of marker loci (e.g. SNPs in candidate genes) to higher density whole genome scans (Varshney, Graner and Sorrels, 2005). These genotypic data, alone or integrated with phenotypic information, may reveal novel aspects of maize gene pools, heterotic groups, haplotype evolution, gene content and parents used in MAS for specific target environments (Fu and Dooner, 2002; Niebur *et al.,* 2004; Crosbie *et al.,* 2006). When properly integrated with phenotypic information and functional genomics, genotypic data of inbred lines should allow breeders to choose parents that, when mated, should provide populations or gene pools enriched for the more desirable combinations of favourable alleles. Such a starting point is a huge advantage in plant breeding because it increases the probability of selecting progeny that are superior to the parents and that approximate a predicted optimum genotype.

MABC is certainly the form of MAS with the most immediate and obvious benefits for maize breeding. MABC is used for three main purposes: selection of transgenes (or of native DNA sequences of the maize genome, whether genes or QTL), elimination of unwanted regions of the donor-parent genome linked to the transgene and selection of unlinked regions of the recurrent-parent genome. With the exception of DNA markers and transgenes, these have been the same goals of backcross breeding since the inception of that method decades ago (Fehr, 1987). Of course, DNA markers enable breeders to identify progeny that contain the desired recombinant chromosomes and donor-parent genome in a more direct manner. Also, MABC facilitates the process of combining more than one transgene in a given inbred line (e.g. "gene or trait stacking or pyramiding"). This reduces the number of generations needed to reach certain stages of a breeding programme and reduces the time needed to produce commercial hybrids for the market. Generic MABC schemes suitable for maize breeding programmes have been described in detail for single genes (Hospital, Chevalet and Mulsant, 1992; Ragot *et al.,* 1995; Frisch, Bohn and Melchinger, 1999a, 1999b; Frisch and Melchinger, 2001a; Hospital, 2001; Ribaut, Jiang and Hoisington, 2002), for QTL (Hospital and Charcosset, 1997; Bouchez *et al.,* 2002) and for gene stacks (Frisch and Melchinger, 2001b). Versions of such schemes have been used in maize breeding programmes in the private sector, often at their continuous nurseries (Ragot *et al.,* 1995). Most recently, MABC has also been adopted as a tool to develop sets of near-isogenic lines (NILs) for genomics research (Peleman and van der Voort, 2003).

Theoretical and simulation studies have been conducted to identify the most efficient MABC protocols. Parameters most commonly studied include the number of individuals genotyped at each generation, the number of markers used, relative selection pressure for recombination around the target locus or global recovery of recurrent parent genome and the number of individuals selected at any generation. Optimal values for each of the above depend on the objective of the MABC approach in terms of quality (required level of recurrent parent genome recovery), speed (fastest possible conversion or set number of generations) and resources (unlimited or limited). While the fastest and highest quality MABC approaches have the most expensive protocols, less intensive approaches can result in significant time savings and quality improvements when compared with conventional backcrossing approaches and at a fraction of the cost of the most expensive MABC protocols.

Frisch, Bohn and Melchinger (1999b) showed that to minimize linkage drag around the target locus (loci), selection of recombination events close to the target locus (loci) should be conducted in the early backcross generations. Frisch and Melchinger (2001a) and Ribaut, Jiang and Hoisington (2002) further demonstrated that minimizing linkage drag around the target locus requires very large numbers of individuals (possibly hundreds) to be genotyped. Hospital and Charcosset (1997) proposed a selection scheme based on selecting a single individual to be backcrossed. By contrast, Frisch and Melchinger (2001a) proposed selecting several individuals and determining the family size of their backcross progeny based on the individuals' genotypes. By using varying rather than constant numbers of individuals or markers at the different backcross generations, it was shown that the number

of marker data points required could be reduced and thus the efficiency of MABC improved (Hospital, Chevalet and Mulsant, 1992; Frisch, Bohn and Melchinger, 1999b). Several studies also showed that using a limited number of markers on non-carrier chromosomes was sufficient to recover more that 95 percent of the recurrent parent genome in three or fewer backcross generations (Hospital, Chevalet and Mulsant, 1992; Visscher, Haley and Thompson, 1996; Servin and Hospital, 2002).

One of the most important lessons from the various theoretical and simulation studies of MABC is that the effects of the different MABC parameters are not independent of each other. With maize, large backcross populations can be generated from a single plant when that plant is used as the male and recurrent parent plants are used as females. Marker systems in maize are also such that very large amounts of marker data can be generated on plants before flowering. Potential MABC protocols are almost endless in maize and identifying the most efficient is only possible on a case-by-case basis. For example, while achieving almost complete recovery of the recurrent parent's genome is necessary for registering backcross-derived lines and hybrids in many European countries, partial recovery might be sufficient to improve the agronomic performance of varieties in developing countries. The optimal MABC protocols for these two strikingly different objectives will be very different. Protocols for the first objective will involve background selection and the use of background markers very close to the target locus (loci). Protocols for the second objective might involve markers for the target locus (loci) only, while relying on successive backcross generations to recover an adequate amount of recurrent parent genome.

Successful examples of MABC in maize include backcrossing of transgenes (Ragot *et al.,* 1995), and QTL for insect resistance (Willcox *et al.,* 2002), flowering maturity (Ragot *et al.,* 2000; Bouchez *et al.,* 2002) and grain yield (Ho, McCouch and Smith, 2002).

Methods of "forward breeding" with DNA markers have also been proposed and implemented by maize breeding programmes. As with the pedigree-based methods of maize breeding favoured by the private sector, many of the "new" methods that utilize genetic data from DNA markers integrated with phenotypic data are essentially a form of recurrent selection, a method that has been in use for several decades (Hallauer and Miranda, 1981). The key advantages of the new versions of recurrent selection are, of course, the availability of genetic data for all progeny at each generation of selection, the integration of genotypic and phenotypic data, and the rapid cycling of generations of selection and information-directed matings at continuous nurseries.

At least two distinct forms of forward breeding with MAS have been proposed: single large-scale MAS (SLS-MAS) (Ribaut and Betrán, 1999) and MARS (Edwards and Johnson, 1994; Lee, 1995; Stam, 1995). A key difference between the methods is that SLS-MAS employs DNA markers at only one generation and attempts to retain genetic variation in regions of the genome unlinked to the DNA markers, while MARS uses markers at each generation, exhausting genetic variation in most regions of the genome. Versions of both SLS-MAS and MARS have been used by breeding programmes in the private sector (Johnson, 2004; Eathington, 2005; Crosbie *et al.,* 2006).

SLS-MAS is of particular interest in pedigree breeding as it consists of screening

and selecting individuals at a few loci at early generations, usually F_2 or F_3, (Eathington, Dudley and Rufener, 1997), using large populations (Ribaut and Betrán, 1999). Individuals displaying homozygous favourable genotypes at the loci of interest are selected and self-pollinated while others are discarded. Self-pollinated progeny of the selected plants then proceed normally through subsequent steps of pedigree breeding. Screening large populations is necessary to ensure that genetic diversity is maintained at regions not under genotypic selection, thereby allowing further phenotypic selection to be conducted. Loci at which marker selection operates can be QTL as described by Ribaut and Betrán (1999). SLS-MAS is thus limited by issues such as the precision of the QTL parameters (position, effect), and relevance of the QTL across environments or gene pools. SLS-MAS can also be conducted for genes, eliminating many of the limitations pertaining to QTL. Although a powerful approach adopted in several species (barley, soybean, sunflower, wheat) to enrich breeding populations at a few loci (Crosbie *et al.,* 2006), SLS-MAS does not appear to have been widely implemented in maize breeding programmes.

MARS targets all traits of importance in a breeding programme and for which genetic information can be obtained. Genetic information is usually obtained from QTL analyses performed on experimental populations and comes in the form of maps of QTL with their corresponding effects. If the QTL mapping analysis is conducted based on a bi-parental population, the sign of the effect at each QTL indicates which of the two parents carried the favourable allele at that QTL. As both parents often contribute favourable alleles, the ideal genotype is a mosaic of chromosomal segments from the two parents. This assumes that the goal is to obtain individuals with as many accumulated favourable alleles as possible, a different goal from that of marker-assisted population improvement as studied elsewhere (Lande and Thompson, 1990; Gimelfarb and Lande, 1994; Gallais, Dillmann and Hospital, 1997; Hospital, Chevalet and Mulsant, 1997; Knapp, 1998; Moreau *et al.,* 1998; Xie and Xu, 1998). Population improvement schemes are generally based on the random mating of selected individuals while the scheme proposed here is based on directed recombination between specific individuals. As reported by Stam (1995), the ideal genotype, defined as the mosaic of favourable chromosomal segments from two parents, will usually never occur in any F_n population of realistic size. It is, however, possible to design a breeding scheme to produce or approach this ideal genotype based on individuals of the experimental population. This breeding scheme could involve several successive generations of crossing individuals (Stam, 1995; Peleman and van der Voort, 2003) and would therefore constitute what is referred to as MARS or genotype construction. This idea can be extended to situations where favourable alleles come from more than two parents (Stam, 1995; Peleman and van der Voort, 2003).

Van Berloo and Stam (1998, 2001) and Charmet *et al.* (1999) developed computer simulations around this idea and assessed the relative merits of marker-assisted genotype construction over phenotypic selec-tion. MARS was simulated in an experimental population where QTL had been mapped. Index (genetic) values were computed for each individual based on its genotypes at QTL-flanking markers (van Berloo and Stam, 1998, 2001). All simulation studies of MARS found that

it was generally superior to phenotypic selection in accumulating favourable alleles in one individual (van Berloo and Stam, 1998, 2001; Charmet *et al.,* 1999). MARS appeared to take better advantage of the genetic diversity present in the populations to which it was applied than phenotypic selection. Simulation research conducted by van Berloo and Stam (2001) showed that MARS was between 3 and almost 20 percent more efficient than phenotypic selection. The advantage of MARS over phenotypic selection was greater when the population under selection was larger or more heterozygous (BC_1s or F_2s vs. RILs, recombinant inbred lines, or DHs, doubled haploids). Although van Berloo and Stam (2001) limited their simulations to populations of up to 200 individuals, their results seem to indicate that the relative advantage of marker-assisted over phenotypic selection would keep increasing as population size increased. The same simulation studies showed that the advantage of marker-assisted over phenotypic selection was larger when dominant QTL were involved in the selection index, or when trait heritability was low in the case of selection for a single trait (van Berloo and Stam, 1998, 2001). These latter observations are of little relevance to most commercial maize breeding programmes, the goal of which is generally the development of inbred lines improved for several traits that will be later combined into superior hybrid varieties. They should, however, increase the appeal of MARS approaches for breeding programmes aimed at developing open-pollinated varieties.

Simulations have also addressed the impact of the amount and quality of QTL information on selection efficiency. Simulation and empirical studies (Beavis, 1994, 1999) showed that QTL mapping experiments based on segregating populations of less than 500 individuals generally revealed only a subset of all QTL affecting the complex traits segregating in these populations. Quantitative trait loci information used in subsequent MARS was therefore necessarily incomplete. Van Berloo and Stam (2001) showed that the relative advantage of MARS over phenotypic selection decreased rapidly when the fraction of the total genotypic variance explained by the QTL included in the selection index decreased. By contrast (van Berloo and Stam, 1998; Charmet *et al.,* 1999), the efficiency of MARS seems to be rather robust to the well-documented (Lee, 1995) uncertainty of QTL genetic locations. The use of genotypic information at markers flanking the QTL possibly explains this observation.

The cost efficiency of MARS was also investigated through simulation (Moreau *et al.,* 2000; Xie and Xu, 1998). When simulating selection for a single trait, Moreau *et al.* (2000) found that, irrespective of the heritability of the trait, MARS was always more cost efficient than phenotypic selection if the cost of genotyping was less than that of evaluating one individual in one plot. When simulating simultaneous selection for multiple traits, Xie and Xu (1998) found that MARS was more cost efficient than phenotypic selection if the cost of genotyping was less than that of phenotyping one individual for all traits. These studies were based on a single generation of MARS. Also, they did not take into consideration any factors besides genotyping and phenotyping costs, although factors influencing the length of a selection cycle or the number of cycles that can be completed in a year can obviously affect the relative economic merits of marker-assisted and phenotypic selection.

In contrast to the abundance of QTL mapping reports, very few accounts of MARS experiments are found in the literature. Moreau, Charcosset and Gallais (2004) compared phenotypic, marker-only, and combined recurrent selection for grain yield and grain moisture at harvest over several cycles and years in maize. Combined selection was based both on phenotypic and marker information while marker-only selection was based on marker information only. Both the marker-only and the combined selection methods constitute MARS approaches. Several combinations of these three methods of selection were applied to the segregating population that served to map the QTL used in marker-based selection indices. Over the six years of the experiment, two cycles of phenotypic selection, two cycles of combined selection, one cycle of combined selection followed by two cycles of marker-only selection, and one cycle of marker-only selection were conducted in parallel. A reassessment of the positions and effects of QTL was conducted after the first cycle for the three schemes containing multiple cycles. All MARS methods were more efficient than phenotypic selection to increase the frequency of favourable alleles at QTL. Nevertheless, Moreau, Charcosset and Gallais (2004) reported no significant difference between marker-assisted and phenotypic selection on the multitrait performance index, although all MARS methods resulted in genetic gain for both grain yield and grain moisture while phenotypic selection resulted in genetic gain for grain yield but an unfavourable evolution of grain moisture. This disappointing result was tentatively explained by the high heritability of the traits, favourable to phenotypic selection, while the percentage of total phenotypic variance explained by the QTL detected for both traits was only about 50 percent. One very encouraging result of this experiment, although Moreau, Charcosset and Gallais (2004) failed to present it as such, was that the first cycle of marker-only selection was as efficient as phenotypic or combined selection in delivering genetic gain. Two conclusions can be drawn from this observation. First, the QTL identified in the initial experimental population were in general not artefacts. Second, selection pressure applied at these QTL, and aimed at fixing alleles identified as favourable, resulted in a change in performance of the selected population in the desired direction when compared with the initial population.

A similar experiment, although based solely on marker-only recurrent selection, was reported by Openshaw and Frascaroli (1997). They conducted MARS in maize simultaneously for four traits, for each of which about ten QTL had been identified. They showed that genetic gain had been achieved in the first cycle of MARS, but that later cycles did not result in any gain. Possible explanations given for these results included uncertainties about QTL parameters (location and effect), interaction effects (epistasis, genetic x environment interaction), and the fact that selection was based on single markers rather than chromosomal segments (Openshaw and Frascaroli, 1997).

Recent communications from several private MARS research programmes (Ragot *et al.*, 2000; Eathington, 2005; Crosbie *et al.*, 2006) revealed large-scale successful applications in maize. Accounts were given of commercial maize hybrids for which at least one of the parental lines was derived through MARS. Eathington (2005) and Crosbie *et al.* (2006) reported that the rates of genetic gain achieved through MARS were about twice those

of phenotypic selection in some reference populations. Marker-only recurrent selection schemes have been implemented for a variety of traits including grain yield and grain moisture (Eathington, 2005), or abiotic stress tolerance (Ragot *et al.,* 2000), and multiple traits are being targeted simultaneously. Selection indices were apparently based on 10 to probably more than 50 loci, these being either QTL identified in the experimental population where MARS was being initiated, QTL identified in other populations, or genes. Marker genotypes are generated for all markers flanking QTL included in the selection indices (Ragot *et al.,* 2000). Plants are genotyped at each cycle and specific combinations of plants are selected for crossing, as proposed by van Berloo and Stam (1998). Several, probably three to four, cycles or MARS are conducted per year using continuous nurseries. In maize, early versions of such schemes have been tested and implemented (Johnson, 2004; Crosbie *et al.,* 2006).

Results reported in these recent communications about private MARS experiments (Ragot *et al.,* 2000; Eathington, 2005) are in sharp contrast to those in earlier publications (Openshaw and Frascaroli, 1997; Moreau, Charcosset and Gallais, 2004). Several factors can explain these discrepancies:

- *Size of the populations submitted to selection at each cycle*. Given reports that increasing population size should result in higher genetic gain through MARS (van Berloo and Stam, 2001,) it is likely that populations submitted to selection in private programmes are rather large, larger than the 160 and 300 individuals reported respectively by Openshaw and Frascaroli (1997) and Moreau, Charcosset and Gallais (2004).
- *Use of flanking versus single markers*. The use of flanking markers for QTL under selection allows better prediction of the genotype at the QTL than when using single markers. When single markers are used, recombination events that occur between the marker and the QTL lead to loss of linkage between the marker and the QTL much faster than when flanking markers are used, thereby rapidly reducing the predictive power of the single marker.
- *Early selection, pre-flowering*. The ability to select plants before flowering ensures optimal mating schemes as the genotypes of plants being selfed or intercrossed are fully known. However, this is not the case when selection cannot take place before flowering and involves intercrossing selfed progenies of selected plants, the genotypes of which might have drifted significantly from those of their genotyped parents.
- *Number of generations per year*. To the authors' knowledge, none of the simulation or experimental studies of MARS has assessed the effects of cycle length on its efficiency despite its direct relationship to the rate of genetic gain. In maize, cycle length can be reduced three- to six-fold when using marker-only recurrent selection compared with phenotypic recurrent selection. Consequently, marker-only recurrent selection will be superior to phenotypic selection as soon as the genetic gain achieved through one cycle of MARS is, respectively, more than a third or a sixth of that achieved through one cycle of phenotypic selection. Private maize breeding programmes have access to off-season nurseries. Furthermore, they have often established efficient continuous nurseries where three to four generations of maize can be grown per year. The use of such nurseries allows them to carry MARS continuously, i.e.

with up to four cycles per year, whereas phenotypic recurrent selection is limited to one cycle per year at most. The impact on the rate of genetic gain of such an implementation of MARS might be very positive even if MARS did not present any advantage over phenotypic selection on a per-cycle basis.

- *Cost of marker data points*. Large private companies have made considerable efforts to reduce both the cost of marker data points and the cost of experimental field plots. The ratio of cost of marker data point to cost of experimental field plot is most likely lower in large private breeding programmes than in most public research laboratories or small private programmes, potentially leading to different views on the economic efficiency of MARS.

Marker-based and phenotypic selection can be mobilized in many different ways, with respect to each other, in marker-assisted breeding schemes. Marker and phenotypic information can be used either simultaneously or sequentially. Selection of parents for breeding populations can be made using marker information alone, phenotypic information alone, or a combination of each. Selection of individuals in a backcross programme can be made on the sole basis of either marker or phenotypic information, or using both. Advancement of individuals in a line development programme can also be made at each generation on the basis of either marker information only, phenotypic information only, or a combination of each. In order to maximize the rate of genetic gain it is likely that MAS breeding schemes such as MABC and MARS will involve generations of marker-only selection conducted at continuous nurseries. The advent of improved methods of producing doubled haploids will certainly further influence the way marker-based and phenotypic selection are mobilized with respect to each other.

In spite of the development of marker-only selection and regardless of the underlying technology and breeding scheme, high-quality phenotyping remains vital and without substitute at several stages; but it may become more focused. Phenotypic evaluation remains the ultimate screen before any cultivar is released. MAS-derived lines and hybrids that meet phenotypic requirements are selected for further evaluation and selection on the basis of their phenotypic value, while those that do not are discarded. Phenotypic evaluation is also critical to establish marker-trait associations or perform the candidate gene validations required to conduct MAS. Here, high quality phenotyping is necessary. Phenotyping protocols will therefore likely be different from those commonly used for phenotypic selection. Experiments may be conducted that involve side-by-side comparisons of different treatments such as water stress or nitrogen fertilization levels to dissect complex traits into their components and facilitate the elucidation of their genetic basis.

Enhancements of such approaches to maize breeding will be based on the incorporation of improved methods of producing doubled haploid inbred lines, information from functional genomics and by learning how to incorporate favourable native genetic variation systematically after MAS has reduced the genetic variation in the original reference populations to unacceptable levels.

ADVANTAGES AND LIMITATIONS OF MAS IN MAIZE BREEDING PROGRAMMES

Advantages of MAS

For private breeding programmes, MAS has offered several attractive features, most

of which are related to time and resource allocations.

MABC clearly provides the information needed to reduce the number of generations of backcrossing, to combine (i.e. "stack") transgenes, "native" genes or QTL into one inbred or hybrid quickly, and to maximize the recovery of the recurrent parent's genome in the backcross-derived progeny. In several private breeding programmes, MABC has enabled the number of backcrossing generations needed to recover 99 percent of the recurrent parent genome to be reduced from six to three, reducing the time needed to develop a converted variety by one year (Crosbie *et al.,* 2006; Ragot *et al.,* 1995). As a line derived by MABC can be made to be very similar to the original non-converted line, most of its attributes, including agronomic performance, can be assumed to be equal or similar to those of the original line. Only limited phenotyping is therefore necessary to verify these assumptions, compared with the extensive multiyear phenotyping required when backcrossing is conducted without markers. One or two years can be saved with MABC during post-conversion phenotyping when compared with conventional backcrossing, resulting in an overall time advantage of MABC over conventional backcrossing of up to three years.

In many situations, the greatest advantages and profits are realized by those who are first to the market with their products. Also, for reasons related to the practices of seed production or legal aspects of crop registration procedures, it may be quite important to be able to produce near-isogenic versions of inbreds and hybrids; MABC provides such ability at a higher probability.

By contrast with MABC, SLS-MAS and MARS do not necessarily decrease the time needed to develop inbred lines. The use of MARS might actually increase it. The advantage of SLS-MAS and MARS resides in their ability to increase the rate of genetic gain (Eathington, 2005), which potentially results in higher performing lines and hybrids than can be developed through phenotypic selection only. Both SLS-MAS and MARS increase the frequency of favourable alleles in the population of selected individuals. The difference between the two approaches is that SLS-MAS operates on few loci while MARS operates on many. When SLS-MAS or MARS are used, the effective size of the population on which selection operates is increased, either directly for SLS-MAS or indirectly through several consecutive generations for MARS when compared with phenotypic pedigree selection. This increase in effective population size permits the application of a greater selection intensity and hence produces a higher genetic gain. SLS-MAS and MARS can also be seen as pre-selection steps if conducted prior to phenotypic selection and therefore improve the chances of evaluating genotypes with a higher frequency of favourable alleles phenotypically because the truly undesirable portion of the population may have been eliminated prior to phenotyping. Phenotypic selection can therefore be conducted with higher selection intensity than would be possible if no pre-selection had taken place, resulting potentially in additional genetic gain.

Alternatively, the resources used for phenotyping can be allocated differently based on whether individuals have been pre-selected or not with MAS. MAS schemes for forward breeding should enable breeding programmes to reallocate or focus resources for phenotypic evaluation in the target environment. For example, if DNA markers are linked to genes for resistance

to a disease or insect then it should be possible initially to select resistant progeny in the absence of the disease or insect by using the DNA data at continuous nurseries. The selected progeny could then be evaluated using relatively more expensive bioassays with the pest(s) in the target environment. This shift in resources is inherent to MARS schemes for complex traits (Edwards and Johnson, 1994; Johnson, 2004; Crosbie *et al.,* 2006). By enriching populations through rapid cycles of MARS at continuous nurseries, breeders should derive a higher frequency of progeny with favourable alleles and haplotypes that are then evaluated in the target environment. Without MARS, resources for evaluation in the target environment would be diluted by the inclusion of too many progeny with an undesirable genetic constitution.

Concerns about reduced genetic diversity among commercial maize hybrids and depletion of genetic diversity in gene pools used in breeding may be partially alleviated by successful implementations of MAS. MABC may revive interest in using essentially untapped maize exotic germplasm as a source of favourable alleles for improvement of elite varieties. Very small and targeted chromosomal segments of exotic origin can be introgressed into elite inbred lines with limited risk of carrying along undesirable characteristics. Such an approach could be beneficial in maize although no accounts of its implementation have been reported despite the many years as reports of its successful use in tomato (Tanksley *et al.,* 1996; Bernacchi *et al.,* 1998a, b; Robert *et al.,* 2001), rice (Xiao *et al.,* 1998), and soybean (Concibio *et al.,* 2003). MARS, in turn, may also contribute to increasing genetic diversity among commercial maize hybrids because, by focusing on selecting specific recombination events, it will result in the development of genuinely new genomic rearrangements. As QTL identified in any experiment represent only a fraction of the loci responsible for the phenotypes of complex traits, one can assume that breeding programmes in different private companies will conduct MARS based on their different genetic models and select for different genomic rearrangements. As a result, hybrids of similar and high performance might be developed that are based on different sets of favourable alleles at different loci, representing distinct "genetic solutions" and contributing to increased genetic diversity in farmers' fields.

An indirect but important advantage of MAS and its underlying information and technology relates to intellectual property. Some maize breeding programmes have created a form of wealth through their collection and knowledge of maize germplasm. Significant investments have been made in maize breeding as exemplified by the billions of United States dollars that were used to purchase a few private programmes between 1995 and 2005. Protecting and maximizing returns on such investments have always been important but are now of greater concern. Information from MAS should be advantageous for addressing issues concerning ownership and derivation of germplasm, relatedness among germplasm and for the formation of some claims in patents and similar documents.

Perhaps one of the greatest advantages of MAS is that, for the first time, maize breeders have the means of learning some of the genetic details about germplasm and the response to selection. Some maize programmes in the private sector have started this process (Niebur *et al.,* 2004). As real functions become associated with the many candidate genes and other DNA sequences,

the opportunities for learning about and understanding the response to selection will increase dramatically. It may then be possible to ameliorate some of the limitations of MAS and truly breed by design.

Limitations of MAS

While not truly an inherent limitation of the methods involved, one unavoidable limitation of MAS is the cost of assembling and integrating the necessary infrastructure and personnel. These can be substantial and beyond the means of many programmes. For such programmes, implementation of MAS could lead to a delusional or unbalanced reallocation of resources from vital activities such as high-quality phenotypic evaluation and selection in the target environment. Currently, only the largest maize breeding programmes in a given market or region have the scale of sales and diversity of products that can justify and support MAS and withstand some of the financial burdens of establishing and replacing components of the system (e.g. changes in the methods and platforms for detecting DNA polymorphisms).

Some inherent limitations to MAS are related to the estimates of QTL position and genetic effects and the rates of false positives and negatives. Confidence intervals for QTL are typically 10–15 cM; a genetic region that should not be a major barrier for implementing MAS although it could become a limitation to achieving genetic gain by preventing the selection of desired recombination events. The advent of association mapping and a growing pool of candidate genes should provide some resources needed to minimize problems related to the estimation of QTL position. The genetic effects of QTL are overestimated for many reasons, some of which are linked to experimental designs for phenotyping or population development while others are inherent to the process of QTL detection (Lee, 1995; Beavis, 1998; Melchinger, Utz and Schön, 1998; Holland, 2004). In addition, genetic effects related to epistasis are either poorly estimated or ignored by programmes in the private sector (Holland, 2001; Crosbie *et al.,* 2006). Such assessments of genetic effects will inflate predictions of genetic gain. The relative merit of MAS will depend on the nature of predictions, actual results and costs of alternative methods.

A possible limitation of MAS with maize is the structure and content of various gene pools. Examples of maize gene pools would include European flint and dent germplasm, United States dents and various heterotic groups within each of these and other larger pools. Surveys with DNA markers have established differences among such groups of germplasm (Smith and Smith, 1992; Niebur *et al.,* 2004). The relatively allele-rich maize gene pools coupled with genetic heterogeneity for many traits will hinder the ability to extrapolate information about genotype-phenotype relationships across gene pools. Such transfer of information is expected to be more successful in relatively homogeneous and less diverse maize gene pools (e.g. sweetcorn or popcorn) and with self-pollinated plant species (Lee, 1995). There have been undocumented reports of a few alleles at QTL that have relatively universal genetic effects across a relatively broad range of maize populations and target environments, but details of such genetic factors have not been publicly disclosed (Crosbie *et al.,* 2006). More resources will need to be devoted to discovering where genetic information cannot be easily extrapolated across gene pools or even populations within a gene pool compared with situations where

it could. Although this should not impact the economic efficiency of MABC or forward breeding, it could affect the overall cost efficiency of MAS.

Finally, the efficacy for MAS in relatively complex populations such as synthetics and open-pollinated varieties (OPVs) has not been investigated. Compared with the bi-allelic populations used in the private sector, such populations are likely to have more than two alleles at a given locus. Also, unlike the simple bi-allelic populations, allele frequency should be an important component of predictions with such populations. Therefore, there should be more genetic effects and interactions to consider when making predictions based on MAS with OPVs and synthetics.

In the future, successful implementation of MAS in maize may lead to more frequent problems related to limited genetic variation. The emphasis of aggressive private sector maize breeding programmes on crosses between elite, related inbred lines to create segregating source populations has led to concerns about the depletion of genetic diversity in such gene pools and the ability to enhance such gene pools with high quality genetic variation (Niebur *et al.*, 2004). Such concerns, which existed prior to the deployment of DNA markers and MAS in maize, are likely to increase as MAS becomes more prevalent. If MAS in forward breeding schemes is as effective as reported, then alleles and haplotypes may approach fixation more rapidly (Crosbie *et al.*, 2006). At that point, breeding programmes will need to repeat the process of calibrating genotype-phenotype relationships in a slightly different array of reference populations to start the next metacycle of MAS (Johnson, 2004).

There is much anticipation for the future of MAS as genic sequences become the marker loci, functional information is discovered for the many candidate genes and gene products are assessed for their potential as useful sources of information in breeding programmes (Varshney, Graner and Sorrels, 2005; Lee, 2006). Certainly, these huge sets of raw data will contribute to progress. Eventually, other sources of genetic variation unrelated to the primary DNA sequence such as DNA methylation will be evaluated for their influence on genotype-phenotype relationships. Currently, epigenetic variation is mostly ignored from that assessment although it is well known that much of the maize genome may be methylated (Kaeppler, 2004) and may be more dynamic than predicted by current genetic models and mechanisms (Fu and Dooner, 2002). Also, the influences of non-coding sequences such as small interfering RNA (siRNA), matrix attachment regions and long-distance regulatory sequences have yet to be considered for their effects on genetic variation and estimates of genetic values used in MAS (Lee, 2006).

Most of the early limitations of MAS, due to the availability or cost of genotypic data, have been overcome. However, the availability or cost of high-quality phenotypic information is becoming one of the major limitations of MAS. During the past 20 years, development of new technologies and automation and miniaturization of laboratory procedures have contributed to reducing the cost of marker data points as well as the time needed to produce them. Large-scale marker laboratories produce marker data points at less than a tenth of the cost of 20 years ago. By contrast, neither cost nor the time required to produce phenotypic data has changed much, if at all, in the same timeframe. As the establishment of marker-trait associations and ultimately

the success of MAS depends on access to high-quality phenotypic data, means will have to be found to decrease the cost of phenotypic information while maintaining or increasing its quality. Alternatively, a greater proportion of budgets needs to be devoted to collecting phenotypic information.

ACHIEVEMENTS OF MAIZE BREEDING PROGRAMMES WITH MAS

In some important ways, maize breeding has gradually changed since the mid 1990s with the advent of genomics. Genetic principles were always an important component of modern maize breeding and now genetic information of various types is seeping into breeding schemes. MAS is the connection between the growing pool of genetic information and actual plant breeding. Establishing and enhancing this connection have been important achievements.

For the simplest breeding scenario, programmes in the private sector have demonstrated that MABC is an effective and routine method to backcross one or more transgenes into established elite inbred lines, the direct parents of commercial hybrids. Hybrids with effective combinations of transgenes have been very successful in the market. Consequently, MAS has accelerated the delivery of some products to the market; an important achievement in competitive economies.

Programmes in the private sector have also demonstrated a sufficient degree of efficacy of MAS methods to secure protection of intellectual property in patents. Methods, ideas and linkage relationships have been included in claims of patents or patent applications related to MAS (e.g. US5 492 547 1996; US6 455 758B1 2002; US2005/0144664A1 2005; WO2005/000006A2 2005; WO2005/014858A2 2005), or the establishment of marker-trait associations (e.g. US5 746 023P 1998; US6 368 806B1 2002; US6 399 855B1 2002). Given the magnitude of the investments made in maize breeding by the private sector, receiving such a legal position may be a valuable achievement for the owner of the patent.

The efficacy of MAS for forward breeding of complex traits has yet to be firmly established. Positive results from calibration studies have been reported, but although accounts of MAS-derived commercial varieties have been made (Eathington, 2005), the impact on actual breeding and the development of new commercial hybrids has not been disclosed to a significant extent (Johnson, 2004; Niebur *et al.*, 2004; Crosbie *et al.*, 2006). At this point in time, it is therefore too early to make a definitive and databased assessment of this aspect of MAS.

The history and cost of the genetic gain achieved through MAS will certainly vary among target environments. In some regions of the world, such as the central United States, maize breeding achieved steady genetic gains in grain yield for several consecutive decades prior to the advent of MAS (Duvick, Smith and Cooper, 2004). Nevertheless, the cost per unit gain has increased as more resources are needed for phenotypic evaluation in more environments (Smith *et al.*, 1999). However, the advent of applied genomics and the discovery of many genes and gene functions, coupled with MAS, could reduce the dependence on costly phenotypic information for breeding. In regions where biotic and abiotic stress factors are more important than in the central Unites States, MAS may be very effective. Ultimately, the value and achievements of MAS will depend on the ecological and socio-economic context of the target environment.

COLLABORATION BETWEEN THE PRIVATE AND PUBLIC SECTORS IN MAS AND MAIZE IMPROVEMENT

The increased investments in maize breeding, expected returns on investment and concerns regarding intellectual property by the private sector have made it more difficult for corporations to collaborate with external parties of any kind. Such factors hinder the exchange of information and material that is common in collaborative projects. Nevertheless, around the world, the private and public sectors still manage to collaborate through various mechanisms and at different levels in the pursuit of maize improvement. Such collaboration involves interactions among multinational corporations, philanthropic foundations, national and subnational governments, universities and individuals. Major categories of collaboration include social programmes and institutions, research and development, and education.

In many regions of the world, private sector maize breeding would not have grown without some critical social programmes and institutions. For example, legislation related to intellectual property, transfer of capital and material, and regulatory approval of biotechnical innovations in maize improvement have been important components of legal systems that have encouraged financial investment in maize breeding. The stability of these systems and the rule of law have contributed to the long-term gains in selection. Also, long-term crop subsidy programmes in some regions have provided an element of security for investments in maize research and development by the private sector (Troyer, 2004; Crosbie *et al.,* 2006). In those same regions, MAS has been deployed initially and on the largest scale for maize breeding.

With respect to research and development, there is a long history of effective collaboration between the public and private sectors in maize breeding. While such interaction continues in the era of MAS, the nature of the collaboration has changed with the growth and development of the breeding programmes in the private sector. Initially, collaboration was absolutely vital for the private sector because breeding programmes in the public sector were important, or the sole, sources of the inbred lines used directly by the private sector to produce commercial hybrids or to source populations from which elite inbreds were derived. Also, the inbred lines from the public sector were usually provided on an unrestricted basis and without payments of royalties or licensing fees. Public breeding programmes continue to develop elite inbred lines, occasionally in collaboration with the private sector (e.g. the Germplasm Enhancement of Maize programme in the United States; Pollak, 2003). However, the direct impact of contemporary public germplasm varies greatly among regions and gradually, in many regions of the world, the private sector has become the primary source of elite maize inbred lines and commercial hybrids.

In addition to germplasm, most or all of the critical concepts, methods and basic technologies have their origins in the public sector (Niebur *et al.,* 2004; Troyer, 2004; Crosbie *et al.,* 2006). The private sector, with its unique ability to concentrate capital through various mechanisms (e.g. profits from products or licence fees, venture capital and financial markets), is in the best position to allocate resources quickly to assess, modify and apply new developments in MAS and ancillary areas of maize improvement across large geographical and political regions of a market zone. As

described in previous sections, cost-effective MAS requires several components of an integrated infrastructure, some features of which have had a relatively high rate of renovation and replacement (e.g. methods of detecting DNA polymorphism), and therefore required substantial financial resources. Competing corporations and the potential for profit provide the necessary motivation for such investments (Troyer, 2004; Crosbie *et al.,* 2006). To the authors' knowledge, such financial mechanisms either do not exist or are limited in the public sector.

Collaboration between the public and private sectors in MAS for maize may be strongest in basic genetics and genome annotation. In order for MAS to reach its full potential, it may be necessary to acquire a much better understanding of gene function and products. For any plant species, only a small percentage of genes and other DNA sequences have a function defined through direct experimentation (Lee, 2006). Discoveries in plant gene function will occur in many laboratories around the world and, ultimately, the development groups in the private sector will have the necessary concentration of resources and sense of purpose to assemble the relatively raw basic information into tools and products from MAS. The maize nuclear genome, with tens of thousands of genes and many other important DNA sequences, is mostly a "black box" with respect to understanding the role of these in mediating phenotypes in response to environmental cues. Such understanding, a potential key to MAS and maize improvement, can only occur through informal and formal collaboration between the public and private sectors investigating a broad array of plant species.

Examples of collaborative research between the public and private sectors relevant to MAS in maize include attempts to select for hybrid yield (Stromberg, Dudley and Rufener, 1994), QTL mapping and selecting of hybrid yield (Eathington, Dudley and Rufener, 1997) and grain quality (Laurie *et al.,* 2004), the development of the *IBM* population of recombinant inbred lines, and mapping genomic regions that include the *vgt1* locus in maize (Lee *et al.,* 2002; Salvi *et al.,* 2002). National collaborative research programmes such as Génoplante in France and GABI in Germany, as well as several projects within the European Commission-sponsored framework programmes, are additional examples of such collaboration. Certainly, other collaborative projects between the public and private sectors have been conducted in maize MAS but their proprietary nature prevents public disclosure.

Future collaborative research activities in maize MAS could assume many forms. In most regions of the world, the private sector has the obvious superiority in terms of infrastructure needed for genotyping, phenotyping and data analysis. These resources are mostly devoted to the direct pursuit of products and profits. That pursuit may also be the greatest disadvantage of the private sector because such a focus limits the attention devoted to many interesting yet seemingly ancillary observations of genotype-phenotype relations in MAS. Some components of that infrastructure could possibly be made accessible to the public sector as "in-kind" contributions to collaborative or service-related projects in regions that are unlikely to emerge as important markets for the private sector or for phenotypes and germplasm that are not of direct interest to the private sector.

Education and training are also important areas in which the public and

private sectors should collaborate. With the advent of MAS, there has been an obvious need for maize breeders in the private sector to become familiar with all aspects of the process, and the public sector has developed several new short courses and training sessions in MAS-related concepts (Niebur *et al.*, 2004; Crosbie *et al.*, 2006).

Such knowledge is now considered a standard component of recent graduate training. However, while new students may have an adequate grasp of the theoretical aspects of MAS, their lack of exposure to the private sector's advanced infrastructure represents a gap in their education. This situation is similar to that of students with a new degree in engineering who join advanced engineering and design groups in other industries: the private sector's capacity to concentrate and focus capital often leads to advanced infrastructure that does not exist in the public sector. In such situations, new students have to navigate a rather steep learning curve before they become productive members of their new group. To reduce the slope of the learning curve, the private sector could provide internships to graduate students or to professors who teach plant breeding courses. It is unlikely that the public sector will have the resources to duplicate or exceed some features of the infrastructure that has been developed for maize MAS in the private sector. Therefore, for some aspects of education, it will be to everyone's benefit to find ways to work together.

PRIVATE SECTOR PERSPECTIVES ON MAS FOR MAIZE IMPROVEMENT

The development of molecular markers in the 1980s provided the first tools to dissect the genetic basis of traits and select individuals based on their predicted genetic value. Back in these early days, the availability of genetic information was a limiting factor. Today's landscape is very different as advances in applied genomics and laboratory technology have provided the tools to generate genetic information for all traits of interest. Gene similarities and synteny across genomes mean that much of the information generated on any plant species has relevance to other plant species. The speed at which genetic information becomes available never ceases to increase. Rather than its availability, it is the ability to handle and utilize genetic information that is becoming the limiting factor for MAS. New and improved information technology and bioinformatics capabilities therefore need to be developed that connect the growing wealth of genetic information with maize breeding programmes where knowledge about the genetic basis of traits and allelic variation at these loci is translated into varieties.

QTL and gene mapping will remain key for the generation and use of genetic information. As sequencing of cereal genomes including maize progresses, physical mapping of cloned genes will become a powerful alternative to statistical approaches. Characterization of allelic diversity at loci of interest can proceed from analyses of bi-parental populations or association studies. An effective alternative is the use of sets of NILs, or introgression line (IL) libraries (Peleman and van der Voort, 2003). As NILs developed around a specific locus differ only by the allele at this locus, and because most traits of agronomic interest in maize are quantitative, phenotypic differences among such NILs are expected to be rather small. High precision phenotyping will not only be required but will be critical for the evaluation of such material (Peleman and van der Voort, 2003). Private corporations have realized the need

for such high precision phenotyping as can be seen from their active recruiting of trait-specific phenotyping scientists often located in targeted areas where the trait of interest can be more easily measured (e.g. positions dedicated to drought tolerance and located in arid regions of the world).

In order to further the implementation of MAS in breeding, increased numbers of marker data points will be required. Private corporations have established or are developing the capacity to produce hundreds of millions of data points per year in service laboratories, distinct from research units. Besides, smaller "biotech" companies are developing technologies that could reduce the cost of each marker data point to a mere few United States cents. Moving to marker systems that are not based on gels is permitting the automation of most laboratory steps. Data points are being produced around the clock with laboratory technicians working in shifts. Here again, private companies are actively recruiting highly qualified technology specialists as well as laboratory managers whose role is more to optimize the running of production plants than dwell on the science. Beyond laboratories, plant handling is becoming a bottleneck to high-throughput protocols. High-throughput facilities have to be established and equipped at continuous nursery sites potentially to handle millions of plants per year.

There is little doubt that the largest private maize breeding programmes are investing very heavily in the implementation of MAS. Unless regulatory issues change dramatically, MABC will remain the preferred means of delivering transgenes to the market. Faster MABC protocols will always represent a potential commercial advantage in an area where competition is fierce and a one-year advantage may mean much on the market. Most recent investments have been directed at implementing MARS in breeding. The size of the investment in this approach seems to suggest that private corporations have more insight into its benefits compared with conventional breeding than has been reported publicly. Genotype-driven breeding should also allow faster development of specialized varieties as the maize market becomes more and more fragmented based on end-use of the harvest: animal feed (silage or grain), ethanol, dry or wet milling. Favourable alleles for traits of interest are likely to be spread across more than two lines therefore requiring the assembly of alleles from many different sources in a single inbred line. Proposals have been made to achieve such goals (Peleman and van der Voort 2003), although software tools to determine the optimal breeding schemes are not yet available to generate these "ideal" genotypes.

Maize breeding is likely to change more in the coming 10 or 20 years than it has over the past 50. Developing new hybrids efficiently now requires integrating data from many sources, sometimes beyond maize, generating high-quality genotypic and phenotypic data needed for the construction of "ideal" genotypes, and finally selecting phenotypically the best individuals from populations of marker-assisted-derived materials. Many stakeholders beyond maize breeders now take an active part in the development of new varieties and therefore breeding will increasingly become the responsibility of groups of individuals with complementary skills than stand-alone breeders. Training of all to understand and challenge the contribution of others will be critical to operating multidisciplinary breeding teams efficiently.

MAS FOR MAIZE IMPROVEMENT IN DEVELOPING COUNTRIES

A rapid analysis of the implementation of MAS in private maize breeding programmes points to three elements as being of particular importance: availability of high-quality phenotypic data, access to low-cost molecular marker data points and access to reliable continuous nurseries. The importance of high-quality phenotypic analyses has been clearly recognized by groups in the private sector (Niebur *et al.*, 2004; Crosbie *et al.*, 2006). Implementation of MAS in maize breeding requires large amounts of marker data points to be generated. Private groups have spent much effort developing technologies and platforms to achieve cost-efficient genotyping. Simultaneously, highly efficient continuous nurseries have been established in tropical environments or local greenhouses.

By contrast, maize breeding for developing countries is rather fragmented. National agricultural research institutions and international centres of the CGIAR such as the International Maize and Wheat Improvement Center (CIMMYT) focus much of their efforts on poor farmers and underserved regions. Private maize breeding programmes are also established in a number of developing countries. Due to the large up-front costs of assembling infrastructure and personnel for genotyping, it is unlikely that individual national marker laboratories could produce data points in a cost-efficient manner. However, regional facilities serving the needs of several national programmes and supported by local laboratories that could process samples (processing samples could be as easy as taking and air-drying them) and provide information in a timely manner, would probably be very sustainable alternatives. Such a regional molecular service laboratory has been established recently in Nairobi, Kenya, in a joint effort by two CGIAR centres, CIMMYT and the International Livestock Research Institute (ILRI) and Kenya's Agricultural Research Institute (KARI), under the Canadian International Development Agency (CIDA)-funded Biosciences eastern and central Africa (BecA) platform, to provide technical access and training for African maize breeders (Delmer, 2005). Such a facility could be an excellent component of a comprehensive maize breeding effort if it is possible to establish and maintain high-quality personnel and facilities for all of the other aspects of maize breeding in key target environments. However, without high-quality capabilities in phenotypic evaluation and selection, molecular laboratories will be worthless. Research projects involving large-scale (transnational) phenotypic evaluations of key genetic material and focused on specific traits (tolerance to biotic or abiotic stresses) should provide genetic information that is both locally relevant and broadly applicable (geographically and in terms of germplasm). Such projects would also spread the cost of phenotyping across all participants but would only be successful with effective transnational coordination.

Private companies running MAS in maize could contribute to its implementation in developing countries in several ways. First, they could make some of their genetic information available, thereby adding to that already available in the public domain. Much information is being generated in the private sector on traits of importance to developing countries such as disease resistance (e.g. grey leaf spot, northern corn leaf blight, *Fusarium* stalk and ear rots), drought tolerance and nitrogen use efficiency. After validation of its relevance

to the germplasm and environments of target areas, this genetic information could be used to select efficiently for specific traits through MAS. Second, private companies could provide access to some of their genotyping or nursery platforms. Genotyping samples for MAS projects in developing countries would not substantially disrupt private companies' own research if conducted in periods of lower activity, and would provide these MAS projects with marker data points for as low a cost as possible. Third and probably most critically, private companies could train scientists from developing countries on the principles, mechanics and logistics of applying and implementing MAS in maize. Scientists in private maize breeding groups have already identified many of the pitfalls and overcome many of the hurdles linked with the implementation of MAS. Transfer of this knowledge to scientists from developing countries would help them immensely to design marker-based breeding schemes adapted to their sets of constraints.

Beyond their contribution to the implementation of MAS in maize in developing countries, private companies could, in very similar ways, contribute to MAS programmes in other species of importance to developing countries but remote from their core interests. Synteny and gene conservation across species should allow some of the maize genetic information to be transferable to other species. Technology platforms and breeding approaches developed for MAS in maize should be good models for other crops and some might be directly usable. Mechanisms or organizations need to be put in place for these transfers of knowledge and technologies to occur from private maize MAS programmes to other crops in developing countries. Private programmes will likely not drive these transfers but might be very willing to contribute or be directly involved in specific projects provided adequate frameworks exist.

Public–private partnerships will need to be established to manage intellectual property issues related to the transfers of information, material or technologies from private companies to developing countries (Naylor *et al.,* 2004). The African Agricultural Technology Foundation (AATF) is one initiative that has been established to deal with such issues. Several private corporations with major investments in MAS in maize have agreed to provide access to germplasm and knowledge for African countries (Naylor *et al.,* 2004; Delmer, 2005).

As with the private sector in Europe and North America, it will be necessary to provide regular and easy access to education and training in maize MAS as the phenotypes and population structures are likely to differ from those encountered by programmes in the private sector in relatively high-input production environments. Also, and in common with the changes in the private sector, some reorganization or restructuring of public sector programmes may be warranted with the advent of more specialized roles for some personnel.

Understanding the genetic basis of traits and cloning and sequencing the underlying genes will not have an impact on poor farmers unless translated into varieties through breeding. Implementing MAS requires significant investments in both people and infrastructures. Some of the most promising marker-based breeding schemes (e.g. MARS), take about as long as conventional breeding schemes to develop improved varieties and therefore require long-term funding commitments. Funding of practical crop improvement has declined for several years, particularly

in the international public sector (Knight, 2003), and as a result investments have favoured research at the expense of practical applications (Naylor *et al.,* 2004). Whether current funding mechanisms based on short-term (two to five years) grants are adequate to allow maize or any other breeding programmes in developing countries to benefit from the much needed advantages of MAS is questionable.

REFERENCES

Basten, C.J., Weir, B.S. & Zeng, Z-B. 1994. Zmap--a QTL cartographer. *In* C. Smith, J.S. Gavora, B. Benkel, J. Chesnais, W. Fairfull, J.P. Gibson, B.W. Kennedy & E.B. Burnside, eds. *Proc. 5th Wrld. Congr. Genet. Appl. Livest. Prodn.,* Vol. *22,* pp. 65–66. Guelph, Ontario, Canada, University of Guelph.

Beavis, W.D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. *Proc. 49th Ann. Corn and Sorghum Ind. Res. Conf.,* pp. 250–266. Washington, DC, USA, American Seed Trade Association.

Beavis, W.D. 1998. QTL analyses: power, precision and accuracy. *In* H. Paterson, ed. *Molecular dissection of complex traits*, pp. 143–162. Boca Raton, FL, USA, CRC Press.

Beavis, W.D. 1999. Integrating the central dogma with plant breeding. *Abstr. Plant & Anim. Genome* 7: 28.

Bernacchi, D., Beck-Bunn, T., Emmatty, D., Eshed, Y., Inai, S., Lopez, J., Petiard, V., Sayama, H., Uhlig, J., Zamir, D. & Tanksley, S. 1998a. Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *Theor. Appl. Genet*. 97: 170–180.

Bernacchi, D., Beck-Bunn, T., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. & Tanksley, S. 1998b. Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theor. Appl. Genet.* 97: 381–397.

Bernardo, R. 1994. Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci.* 34: 20–25.

Bernardo, R. 2001. What if we knew all the genes for a quantitative trait in hybrid crops ? *Crop Sci.* 41: 1–4.

Bernardo, R. 2002. *Breeding for quantitative traits in plants*. Woodbury, MN, USA, Stemma Press.

Botstein, D., White, R.L., Skolnick, M. & Davis, R.W. 1980. Construction of genetic linkage map in human using restriction fragment length polymorphisms. *Am. J. Human Genet*. 32: 314–331.

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.

Burr, B., Burr, F.A., Thompson, K.H., Albertsen, M.C. & Stuber, C.W. 1988. Gene mapping with recombinant inbreds in maize. *Genetics* 118: 519–526.

Castiglioni, P., Ajmone-Marsan, P., van Wijk, R. & Motto, M. 1999. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. *Theor. Appl. Genet*. 99: 425–431.

Charmet, G., Robert, N., Perretant, M.R., Gay, G., Sourdille, P., Groos, C., Bernard, S. & Bernard, M. 1999. Marker-assisted recurrent selection for cumulating additive and interactive QTL's in recombinant inbred lines. *Theor. Appl. Genet*. 99: 1143–1148.

Ching, A., Caldwell, K.S., Jung, M., Dolan, M., Smith, O.S., Tingey, S., Morgante, M. & Rafalski, A. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics* 3: 19.

Coe, E., Hancock, D., Kowalewski, S. & Polacco, M. 1995. Gene list and working maps. *Maize Genet. Coop. Newsl.* 69: 191–267.

Concibio, V.C., La Vallee, B., McLaird, P., Pineda, N., Meyer, J., Hummer, L., Yang, J., Wu, K. & Delannay, X. 2003. Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor. Appl. Genet.* 106: 575–582.

Crosbie, T.M., Eathington, S.R., Johnson, G.R., Edwards, M., Reiter, R., Stark, S., Mohanty, R.G., Oyervides, M., Buehler, R.E., Walker, A.K., Dobert, R., Delannay, X., Pershing, J.C., Hall, M.A. & Lamkey, K.R. 2006. Plant breeding: past, present, and future. *In* K.R. Lamkey & M. Lee, eds. *Plant Breeding: The Arnel R. Hallauer Internat. Symp.* Ames, IA, USA., Blackwell Publishing.

Davis, G.L., McMullen, M.D., Baysdorfer, C., Musket, T., Grant, D., Staebell, M., Xu, G., Polacco, M., Koster, L., Melia-Hancock, S., Houchins, K., Chao, S. & Coe Jr., E.H. 1999. A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. *Genetics* 152: 1137–1172.

Delmer, D.P. 2005. Agriculture in the developing world: connecting innovations in plant research to downstream applications. *Proc. Natl. Acad. Sci. USA* 102**:** 15739–15746.

Duvick, D.N., Smith, J.S.C. & Cooper, M. 2004. Long-term selection in a commercial hybrid maize breeding program. *In* J. Janick, ed. *Plant Breeding Rev.* 24: 109–152.

Eathington, S.R. 2005. Practical applications of molecular technology in the development of commercial maize hybrids*. Proc. 60th Ann. Corn and Sorghum Seed Res. Conf.* Washington, D.C., USA, American Seed Trade Association.

Eathington, S.R., Dudley, J.W. & Rufener, G.K. 1997. Usefulness of marker-QTL associations in early generation selection. *Crop Sci.* 37: 1686–1693.

Edwards, M. & Johnson, L. 1994. RFLPs for rapid recurrent selection. *Proc. Symp. Analysis of Molecular Marker Data,* pp. 33–40. Corvallis, OR, USA, American Society of Horticultural Science and Crop Science Society of America.

Edwards, M.D. & Page, N.J. 1994. Evaluation of marker-assisted selection through computer simulation. *Theor. Appl. Genet.* 88: 376–382.

Fehr, W.R. 1987. *Principles of cultivar development*. New York, USA, Macmillan.

Frisch, M. & Melchinger, A.E. 2001a. Marker-assisted backcrossing for introgression of a recessive gene. *Crop Sci.* 41: 1485–1494.

Frisch, M. & Melchinger, A.E. 2001b. Marker-assisted backcrossing for simultaneous introgression of two genes. *Crop Sci.* 41: 1716–1725.

Frisch, M., Bohn, M. & Melchinger, A.E. 1999a. Minimal sample size and optimal positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. *Crop Sci.* 39: 967–975.

Frisch, M., Bohn, M. & Melchinger, A.E. 1999b.Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Sci.* 39: 1295–1301.

Fu, H. & Dooner, H.K. 2002. Intraspecific violation of genetic co-linearity and its implications in maize. *Proc. Natl. Acad. Sci. USA* 99: 9573–9578.

Gallais, A., Dillmann, C. & Hospital, F. 1997. An analytical approach of marker assisted selection with selection on markers only. *In* R. Krajewski & Z. Kaczmarek, eds. *Advances in biometrical genetics. Proc.10th. Mtg. EUCARPIA Section Biometrics in Plant Breeding,* pp. 111–116. Poznan, Institute of Plant Genetics, Polish Academy of Sciences.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. 2004. *Bayesian data analysis*, 2nd edition. Boca Raton, FL, USA, Chapman and Hall/CRC.

Gimelfarb, A. & Lande, R. 1994. Simulation of marker-assisted selection in hybrid populations. *Genet. Res*. 63: 39–47.

Goodman, M.M. & Carson, M.L. 2000. Reality vs. myth: corn breeding, exotics, and genetic engineering. *Proc. 55th Ann. Corn and Sorghum Res. Conf.,* pp. 149–172. Washington, D.C., USA, American Seed Trade Association.

Goodman, M.M., Stuber, C.W., Newton, K. & Weissinger, H.H. 1980. Linkage relationships of 19 enzyme loci in maize. *Genetics 96*: 697–710.

Hallauer, A.R. & Miranda, J.B. 1981. *Quantitative genetics in maize breeding*. Ames, IA, USA, Iowa State University Press.

Helentjaris, T., Slocum, M., Wright, S., Schasfer, A. & Nienhuis, J. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theor. Appl. Genet*. 72: 761–769.

Ho, J.C., McCouch, S.R. & Smith, M.E. 2002. Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. *Theor. Appl. Genet.* 105: 440–448.

Hoisington, D. 1989. Working linkage maps. *Maize Genet. Coop. Newsl.* 63: 141–151.

Holland, J.B. 2001. Epistasis and plant breeding. *In* J. Janick, ed. *Plant Breeding Rev.* 21: 29–32.

Holland, J.B. 2004. Implementation of molecular markers for quantitative traits in breeding programs – challenges and opportunities. *In* T. Fischer, N. Turner, J. Angus, L. McIntyre, M. Robertson, A. Borrell & D. Lloyd, eds. *New directions for a diverse planet. Proc. 4th Internat. Crop Sci. Congr.,* Brisbane, Australia (available at www.cropscience.org.au/icsc2004).

Hospital, F. 2001. Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. *Genetics* 158: 1363–1379.

Hospital, F. & Charcosset, A. 1997. Marker-assisted introgression of quantitative trait loci. *Genetics* 147: 1469–1485.

Hospital, F., Chevalet, C. & Mulsant, P. 1992. Using markers in gene introgression breeding programs. *Genetics* 132: 1199–1210.

Jansen, R.C. 1993. Interval mapping of multiple quantitative trait loci by using molecular markers. *Theor. Appl. Genet*. 85: 252–260.

Jansen, R.C., Jannink, J-L. & Beavis, W.D. 2003. Mapping quantitative trait loci in plant breeding populations: use of parental haplotype sharing. *Crop Sci.* 43: 829–834.

Johnson, G.R. 2004. Marker-assisted selection. *In* J. Janick, ed. *Plant Breeding Rev.* 24: 293–310.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., van de Wiel, C., Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volkaert, G., Rueda, J., Linacero, R., Vazquez, A. & Karp, A. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol. Breed*. 3: 381–390.

Jourjon, M-F., Jasson, S., Marcel, J., Ngom, B. & Mangin, B. 2005. MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics* 21: 128–130.

Kao, C-H., Zeng, Z-B. & Teasdale, R. 1999. Multiple interval mapping for quantitative trait loci. *Genetics* 152: 1203–1216.

Kaeppler, S. 2004. Biotechnology: new horizons. *In* C.W. Smith, J. Betrán, & E.C.A. Runge, eds. *Corn: origin, history, technology, and production*, pp. 399–425. Hoboken, NJ, USA, John Wiley & Sons.

Knapp, S.J. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. *Crop Sci.* 38: 1164–1174.

Knight, J. 2003. Crop improvement: a dying breed. *Nature* 421: 568–570.

Lande, R. & Thompson, R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124: 743–756.

Lander, E.S. & Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185–199.

Laurie, C.C., Chasalow, S.D., LeDeaux, J.R., McCarroll, R., Bush, D., Hauge, B., Lai, C., Clark, D., Rocheford, T.R. &. Dudley, J.W. 2004. The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics* 168: 2141–2155.

Lee, M. 1995. DNA markers in plant breeding programs. *Adv. Agron.* 55: 265–344.

Lee, M. 2006. The phenotypic and genotypic eras of plant breeding. *In* K.R. Lamkey & M. Lee, eds. *Plant Breeding: The Arnel R. Hallauer Internat. Symp.* Ames, IA, USA, Blackwell Publishing.

Lee, M., Sharopova, N., Beavis, W.D., Grant, D., Katt, M., Blair, D. & Hallauer, A.R. 2002. Expanding the genetic map of maize with the intermated B73 x Mo17 (*IBM*) population. *Plant Mol. Biol.* 48: 453–461.

Lindblad-Toh, K., Winchester, E., Daly, M., Wang, D., Hirschhorn, J.N., Laviolette, J.P., Ardie, K., Reich, D.E., Robinson, E., Sklar, P., Shah, N., Thomas, D., Fan, J.B., Gingeras, T., Warrington, J., Patil, N., Hudson, T.J. & Lander, E.S. 2000. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nature Genet.* 24: 381–386.

Manly, K.F. & Olson, J.M. 1999. Overview of QTL mapping software and introduction to Map Manager QT. *Mamm. Genome* 10: 327–334.

Mazur, B.J. 1995. Commercializing the products of plant biotechnology. *Trends Biotech*. 13: 319–323.

Melchinger, A.E., Utz, H.F. & Schön, C.C. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* 149: 383–403.

Mohan, M., Nair, S., Krishna, T.G., Yano, M., Bhatia, C.R. & Sasaki, T. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breed*. 3: 87–103.

Moreau, L., Charcosset, A. & Gallais, A. 2004. Experimental evaluation of several cycles of marker-assisted selection in maize. *Euphytica* 137: 111–118.

Moreau, L., Charcosset, A., Hospital, F. & Gallais, A. 1998. Marker-assisted selection efficiency in populations of finite size. *Genetics* 148: 1353–1365.

Moreau, L., Lemarié, S., Charcosset, A. & Gallais, A. 2000. Economic efficiency of one cycle of marker-assisted selection. *Crop Sci.* 40: 329–337.

Morris, M., Dreher, K., Ribaut, J-M. & Khairallah, M. 2003. Money matters (II): costs of maize inbred line conversion schemes at CIMMYT using conventional and marker-assisted selection. *Mol. Breed*. 11: 235–247.

Naylor, R.L., Falcon, W.P., Goodman, R.M., Jahn, M.M., Sengooba, T., Tefera, H. & Nelson, R.J. 2004. Biotechnology in the developing world: a case for increased investments in orphan crops. *Food Policy* 29: 15–44.

Niebur, W.S., Rafalski, J.A., Smith, O.S. & Cooper, M. 2004. Applications of genomics technologies to enhance rate of genetic progress for yield of maize within a commercial breeding program. *In* T. Fischer, N. Turner, J. Angus, L. McIntyre, M. Robertson, A. Borrell & D. Lloyd, eds. *New directions for a diverse planet. Proc. 4th Internat. Crop Sci. Congr.,* Brisbane, Australia. (available at www.cropscience.org.au/icsc2004).

Openshaw, S. & Frascaroli, E. 1997. QTL detection and marker-assisted selection for complex traits in maize. *Proc. 52nd Ann. Corn and Sorghum Res. Conf.,* pp. 44–53. Washington, DC, American Seed Trade Association.

Parisseaux, B. & Bernardo, R. 2004. *In silico* mapping of quantitative trait loci in maize. *Theor. Appl. Genet*. 109: 508–514.

Peleman, J.D. & van der Voort, J.R. 2003. Breeding by design. *Trends Plant Sci.* 7: 330–334.

Penner, G.A., Bush, A., Wise, R., Kim, W., Domier L., Kasha, K., Laroche, A., Scoles, G., Molner, S.J. & Fedak, G. 1993. Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR Meth. Appl*. 2: 341–345.

Pollak, L.M. 2003. The history and success of the public-private project on germplasm enhancement of maize (GEM). *Adv. Agron.* 78: 45–87.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* 5: 94–100.

Ragot, M. & Hoisington, D.A. 1993. Molecular markers for plant breeding: comparisons of RFLP and RAPD genotyping costs. *Theor. Appl. Genet*. 86: 975–984.

Ragot M., Biasiolli, M., Delbut, M.F., Dell'Orco, A., Malgarini L., Thevenin, P., Vernoy, J., Vivant, J., Zimmermann, R. & Gay, G. 1995. Marker-assisted backcrossing: a practical example. *In* A. Bervillé & M. Tersac, eds. *Les Colloques, No. 72, Techniques et utilisations des marqueurs moléculaires,* pp. 45–56. Paris, France, INRA.

Ragot, M., Gay, G., Muller, J-P. & Durovray, J. 2000. Efficient selection for the adaptation to the environment through QTL mapping and manipulation in maize. *In* J-M. Ribaut & D. Poland, eds. *Molecular approaches for the genetic improvement of cereals for stable production in water-limited environments*, pp. 128–130. México, D.F., CIMMYT.

Ribaut, J-M. & Betrán, J. 1999. Single large-scale marker-assisted selection (SLS-MAS). *Mol. Breed*. 5: 531–541.

Ribaut, J-M., Jiang, C. & Hoisington, D. 2002. Efficiency of a gene introgression experiment by backcrossing. *Crop Sci.* 42: 557–565.

Ribaut, J-M., Edmeades, G., Perotti, E. & Hoisington, D. 2000. QTL analyses, MAS results, and perspectives for drought-tolerance improvement in tropical maize. *In* J-M. Ribaut & D. Poland, eds. *Molecular approaches for the genetic improvement of cereals for stable production in water-limited environments,* pp. 131–136. México, D.F., CIMMYT.

Robert, V.J.M., West, M.A.L., Inai, S., Caines, A., Arntzen, L., Smith, J.K. & St-Clair, D.A. 2001. Marker-assisted introgression of blackmold resistance QTL alleles from wild *Lycopersicon chesmanii* to cultivated tomato (*L. esculentum*) and evaluation of QTL phenotypic effects. *Mol. Breed*. 8: 217–233.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S., Higuchi, R.H., Horn, G.T., Mullis, K.B. & Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–491.

Salvi, S., Tuberosa, R., Chiapparino, E., Maccaferri, M., Veillet, S., van Beuningen, L., Isaac, P., Edwards, K. & Phillips, R.L. 2002. Towards positional cloning of Vgt1, a QTL controlling the transition from the vegetative to the reproductive phase in maize. *Plant Mol. Biol.* 48: 601–613.

Sax, K. 1923. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8: 552–560.

Senior, M.L. & Heun, M. 1993. Mapping maize microsatellites and polymerase chain reaction confir-

mation of the target repeats using a CT primer. *Genome* 36: 884–889.

Senior, M.L., Chin, E.C.L., Lee, M., Smith, J.S.C. & Stuber, C.W. 1996. Simple sequence repeat markers developed from maize sequences found in the GENEBANK database: Map construction. *Crop Sci.* 36: 1676–1683.

Servin, B. & Hospital, F. 2002. Optimal positioning of markers to control genetic background in marker-assisted backcrossing. *J. Hered.* 93: 214–217.

Shattuck-Eidens, D.M., Bell, R.N., Neuhausen, S. & Helentjaris, T. 1990. DNA sequence variation within maize and melon: Observations from polymerase chain reaction and direct sequencing. *Genetics* 126: 207–217.

Skroch, P. & Nienhuis, J. 1995. Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. *Theor. Appl. Genet*. 91: 1086–1091.

Smith, J.S.C. &. Smith, O.S. 1992. Fingerprinting crop varieties. *Adv. Agron*. 47: 85–140.

Smith, O.S., Hoard, K., Shaw, F. & Shaw, R. 1999. Prediction of single-cross performance. *In* J. G. Coors & S. Pandey, eds. *The genetics and exploitation of heterosis in crops,* pp. 277–285. Madison, WI, USA, American Society of Agronomy.

Sprague, G.F. &. Dudley, J.W. 1988. *Corn and corn improvement*, 3rd edition. Madison, Wisconsin, USA, American Society of Agronomy.

Stam, P. 1995. Marker-assisted breeding. *In* J.W. Van Ooijen & J. Jansen, eds. *Biometrics in plant breeding: applications of molecular markers*. *Proc. 9th Mtg. EUCARPIA Section on Biometrics in Plant Breeding,* pp. 32–44. Wageningen, Netherlands, CPRO-DLO.

Stromberg, L.D., Dudley, J.W. & Rufener, G.K. 1994. Comparing conventional early generation selection with molecular marker assisted selection in maize. *Crop Sci.* 34: 1221–1225.

Stuber, C.W. 1992. Biochemical and molecular markers in plant breeding. *In* J. Janick, ed. *Plant Breeding Rev.* pp 37–61.

Stuber, C.W. & Moll, R.H. 1972. Frequency changes of isozyme alleles in a selection experiment for grain yield in maize (*Zea mays* L.). *Crop Sci.* 12: 337–340.

Tanksley, S.D., Young, N.D., Patterson, A.H. & Bonierbale, M.W. 1989. RFLP mapping in plant breeding: New tools for an old science. *Biotech.* 7: 257–264.

Tanksley, S.D., Grandillo, S., Fulton, T.M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. & Beck-Bunn, T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theor. Appl. Genet.* 92: 213–224.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA. *Nucl. Acids Res.* 17: 6463–6471.

Thornsberry, J.M., Goodman, M.M., Doebley, J., Kresovich, S., Nielsen, D. & Buckler, E.S. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genet.* 28: 286–289.

Troyer, A.F. 2004. Persistent and popular germplasm in seventy centuries of corn evolution. *In* C.W. Smith, J. Betrán & E.C.A. Runge, eds. *Corn: origin, history, technology, and production,* pp. 133–232. Hoboken, NJ, USA, John Wiley & Sons.

Utz, H.F. & Melchinger, A.E. 1996. PLABQTL: a program for composite interval mapping of QTL. *J. Quant. Trait Loci* 2: 1.

van Berloo, R. & Stam, P. 1998. Marker-assisted selection in autogamous RIL populations: a simulation study. *Theor. Appl. Genet*. 96: 147–154.

van Berloo, R. & Stam, P. 2001. Simultaneous marker-assisted selection for multiple traits in autogamous crops. *Theor. Appl. Genet*. 102: 1107–1112.

van Ooijen, J.W. & Maliepaard, C. 1996. MapQTL version 3.0: software for the calculation of QTL positions on genetic maps. *Plant Genome IV,* San Diego, CA, USA (available at www.intl-pag.org/4/abstracts/p316).

Varshney, R.K., Graner, A. & Sorrels, M.E. 2005. Genomics-assisted breeding for crop improvement. *Trends Plant Sci*. 10: 621–630.

Visscher, P.M., Haley, C.S. & Thompson, R. 1996. Marker-assisted introgression in backcross breeding programs. *Genetics* 144: 1923–1932.

Vos P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucl. Acids Res*. 23: 4407–4414.

Welsh, J. & McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucl. Acids Res.* 18: 7213–7218.

Willcox, M.C., Kairallah, M.M., Bergvinson, D., Crossa, J., Deutsch, A., Edmeades, G.O., González-de-León, D., Jiang, C., Jewell, D.C., Mihm, A J., Williams, W.P. & Hoisington, D. 2002. Selection for resistance to southwestern corn borer using marker-assisted and conventional backcrossing. *Crop Sci*. 42: 1516–1528.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. & Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res*. 18: 6531–6535.

Xiao, J., Li, J., Grandillo, S., Nag Ahn, S., Yuan, L., Tanksley, S.D. & McCouch, S.R. 1998. Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150: 899–909.

Xie, C. & Xu, S. 1998. Efficiency of multistage marker-assisted selection in the improvement of multiple quantitative traits. *Heredity* 80: 489–498.

Young, N.D. 1999. A cautiously optimistic view for marker-assisted breeding. *Mol. Breed.* 5: 505–510.

Yousef, G.G. & Juvik, J.A. 2001. Comparison of phenotypic and marker-assisted selection form quantitative traits in sweet corn. *Crop Sci*. 41: 645–655.

Zeng, Z-B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc. Natl. Acad. Sci. USA* 90: 10972–10976.

Zeng, Z-B. 1994. Precision mapping of quantitative trait loci. *Genetics* 136: 1457–1468.

CHAPTER 9

Molecular marker-assisted selection for resistance to pathogens in tomato

Amalia Barone and Luigi Frusciante

SUMMARY

Since the 1980s, the use of molecular markers has been suggested to improve the efficiency of releasing resistant varieties, thus overcoming difficulties met with classical breeding. For tomato, a high-density molecular map is available in which more than 40 resistance genes are localized. Markers linked to these genes can be used to speed up gene transfer and pyramiding. Suitable PCR markers targeting resistance genes were constructed directly on the sequences of resistance genes or on restriction fragment length polymorphisms (RFLPs) tightly linked to them, and used to select resistant genotypes in backcross schemes. In some cases, the BC_5 generation was reached, and genotypes that cumulated two homozygous resistant genes were also obtained. These results supported the feasibility of using marker-assisted selection (MAS) in tomato and reinforcing the potential of this approach for other genes, which is today also driven by the development of new techniques and increasing knowledge about the tomato genome.

INTRODUCTION

Tomato (*Solanum lycopersicum*, formerly *Lycopersicon esculentum*) is one of the most widely grown vegetable crops in the world. It is used as a fresh vegetable and can also be processed and canned as a paste, juice sauce, powder or as a whole. World volume has increased approximately 10 percent since 1985, reflecting a substantial increase in dietary use of the tomato. Nutritionally, tomato is a significant source of vitamins A and C. Furthermore, recent studies have shown the importance of lycopene, a major component of red tomatoes, which has antioxidant properties and may help to protect against cancer and heart disease (Rao and Agarwal, 2000).

One of the main constraints to tomato cultivation is damage caused by pathogens, including viruses, bacteria, nematodes and fungi, which cause severe losses in production. The control of pathogen spread mainly involves three strategies: husbandry practices, application of agrochemicals and use of resistant varieties. Husbandry techniques generally help to restrict the spread of pathogens and their vectors as well as to keep plants healthy, thus allowing pathogen attack to be limited. Chemical control gives good results for some pathogens, but poor results against others, such as bacteria, and has practically no effect on viruses. Moreover, reducing chemical treatments lowers the health risks to farmers and consumers. Therefore, in order to achieve sustainable agriculture and obtain high-quality, safe and healthy products, the use of resistant varieties is one of the principal tools to reduce pathogen damage.

Since the early part of the twentieth century, breeding for disease resistance has been a major method for controlling plant disease. Varieties that are resistant or tolerant to one or a number of specific pathogens are already available for many crops, and hybrids with multiple resistance to several pathogens exist and are currently used in vegetable production. In tomato, genetic control of pathogens is a very useful practice with most resistance being monogenic and dominant. Various sources of resistance have been used in traditional breeding programmes, and resistant breeding lines, varieties and F_1 hybrids have been developed with varying stability and levels of expression (Table 1) (Laterrot, 1996; Gardner and Shoemaker, 1999; Scott, 2005).

TABLE 1
List of pathogen resistances present in tomato breeding lines, varieties and F_1 hybrids obtained through conventional breeding

Virus
Beet curly top virus (BCTV)
Tobacco mosaic virus (TMV)
Tomato mosaic virus (ToMV)
Tomato yellow leaf curl virus (TYLCV)
Tomato spotted wilt virus (TSWV)
Bacteria
Corynebacterium michiganense
Pseudomonas solanacearum
Pseudomonas syringae pv. *tomato*
Nematodes
Meloidogyne spp.
Fungi
Alternaria alternata f. sp. *lycopersici*
Alternaria solani
Cladosporium fulvum
Fusarium oxysporum f. sp. *lycopersici*
Fusarium oxysporum f. sp. *radicis-lycopersici*
Phytophthora infestans
Pyrenochaeta lycopersici
Stemphylium solani
Verticillium dahliae

Modified from Laterrot (1996) and updated as reported in the text.

MARKER-ASSISTED BREEDING FOR PATHOGEN RESISTANCE

Although conventional plant breeding has had a significant impact on improving tomato for resistance to important diseases, the time-consuming process of

TABLE 2
Resistance genes mapped on the tomato molecular map

Pathogen	Gene[1]	Chromosomal location	Reference
Virus			
Alfalfa mosaic virus (AMV)	*Am*	6	Parrella *et al*., 2004
Cucumber mosaic virus (CMV)	*Cmr*	12	Stamova and Chetelat, 2000
Potato virus Y (PVY)	*pot-1*	3	Parrella *et al*., 2002
Tomato mottle virus (ToMoV)	2 genes	6	Griffiths and Scott, 2001
Tobacco mosaic virus (TMV)	*Tm-1, Tm2a*	2, 9	Young and Tanksley, 1988; Levesque *et al*., 1990
Tomato spotted wilt virus (TSWV)	*Sw5*	9	Stevens, Lamb & Rhoads, 1995
Tomato yellow leaf curl virus (TYLCV)	*Ty-1* (Q), *Ty-2*	6, 11	Zamir *et al*., 1994; Chagué *et a*l., 1997; Hanson *et al*., 2000
Bacteria			
Clavibacter michiganensis	*Cm1.1- Cm 10.1* (Q)	1, 6, 7, 8, 9, 10	Sandbrink *et al*., 1995
	QTLs	5, 7, 9	van Heusden *et al*., 1999
	Rcm2.0 (Q), *Rcm5.1* (Q)	2, 5	Kabelka, Franchino & Francio, 2002
Pseudomonas syringae pv. *tomato*	*Prf* *Pto*	6 6	Salmeron *et a*l., 1996 Martin *et al*., 1993
Ralstonia solanacearum	*Bw 1, Bw 3, Bw 4, Bw 5* (Q)	6, 10, 4, 6	Danesh *et al*., 2004; Thoquet *et al*., 1996
Xanthomonas campestris pv *vesicatoria*	*Bs4*	5	Ballvora *et al*., 2001
	rx-1, rx-2, rx-3	1	Yu *et al*., 1995
Nematodes			
Globodera rostochiensis	*Hero*	4	Ganal *et al*., 1995
Meloidogyne spp.	*Mi, Mi-3, Mi-9*	6, 12, 6	Williamson *et al*., 1994; Yaghoobi *et al*., 1995; Ammiraju *et al*., 2003
Fungi			
Alternaria alternata f. sp. *lycopersici*	*Asc*	3	van der Biezen, Glagotlkaya & Overduin, 1995
	QTLs 2a, 2c, 3, 9, 12	2, 3, 9, 12	Robert *et al*., 2001
Alternaria solani	*EBR*-QTLs	All	Foolad *et al*., 2002; Zhang *et al*., 2003
Cladosporium fulvum	*Cf-1, Cf-2, Cf-4, Cf-5, Cf-9*	1, 6, 1, 6, 1	Balint-Kurti *et al*., 1994; Jones *et al*., 1993
Fusarium oxysporum f. sp. *radicis- lycopersici*	*Frl*	9	Vakalounakis *et al*., 1997
Fusarium oxysporum f. sp. *lycopersici*	*I1, I2, I3*	7, 11, 7	Bournival, Vallejos and Scott, 1990; Sarfatti *et al*., 1991; Tanksley and Costello, 1991; Ori *et al*., 1997
Leveillula taurica	*Lv*	12	Chunwongse *et al*., 1994
Oidium lycopersicon	*Ol-1, ol-2, Ol-3, Ol-4*	6, 4, 6, 6	Huang *et al*., 2000; Bai *et al*., 2004; De Giovanni *et al*., 2004
	Ol-qtl1, Ol-qtl2, Ol-qtl3	6, 12	Bai *et al*., 2003
Phytophthora infestans	lb1-lb12 (Q)	All	Brouwer, Jones and St. Clair, 2004
	Ph-1, Ph-2, Ph-3	7, 10, 9	Moreau *et al*., 1998; Chunwongse *et al*., 2002
Pyrenochaeta lycopersici	*py-1*	3	Doganlar *et al.,* 1998
Stemphylium spp.	*Sm*	11	Behare *et al*., 1991
Verticillium dahliae	*Ve1, Ve2*	9	Diwan *et al*., 1999; Kawchuck *et al*., 2001

[1] Q in parenthesis, QTL and qtl indicate quantitative trait loci. Recessive resistance genes are reported with small letters.

making crosses and backcrosses, and the selection of the desired resistant progeny, make it difficult to respond adequately to the evolution of new virulent pathogens. Moreover, several interesting resistances are difficult to use because the diagnostic tests often cannot be developed due to the challenge posed by inoculum production and maintenance. In addition, where symptoms are detectable only on adult plants and/or fruits, diagnostic tests can be particularly expensive and difficult to perform.

Since the 1980s, the use of molecular markers has been suggested as a tool for breeding many crops, including tomato. In the last two decades, molecular markers have been employed to map and tag major genes and quantitative trait loci (QTL) involved in monogenic and polygenic resistance control, known respectively as vertical and horizontal resistance. To date, more than 40 genes (including many single genes and QTL) that confer resistance to all major classes of plant pathogens have been mapped on the tomato molecular map (Table 2) and/or cloned from *Solanaceous* species, as reported by Grube, Radwanski and Jahn (2000). Since then, other resistance genes together with resistance gene analogues (RGAs), which are structurally related sequences based on the protein domain shared among cloned R genes (Leister *et al*., 1996), have been added to the map. A molecular linkage map of tomato based on RGAs has also been constructed in which 29 RGAs were located on nine of the 12 tomato chromosomes (Foolad *et al*., 2002; Zhang *et al.,* 2002). Several RGA loci were found in clusters and their locations coincided with those of several known tomato R genes or QTL. This map provides a basis for further identifying and mapping genes and QTL for disease resistance and will be useful for MAS.

In fact, independently of the type of marker used for selection, by making it possible to follow the gene under selection through generations rather than waiting for phenotypic expression of the resistance gene, markers tightly linked to resistance genes can greatly aid disease resistance programmes. In particular, genetic mapping of disease resistance genes has greatly improved the efficiency of plant breeding and also led to a better understanding of the molecular basis of resistance.

DNA marker technology has been used in commercial plant breeding programmes since the early 1990s, and has proved helpful for the rapid and efficient transfer of useful traits into agronomically desirable varieties and hybrids (Tanksley *et al*., 1989; Lefebvre and Chèvre, 1995). Markers linked to disease resistance loci can now be used for MAS programmes, thus also allowing several resistance genes to be cumulated in the same genotype ("pyramiding" of resistance genes), and they may be also useful for cloning and sequencing the genes. In tomato, several resistance genes have been sequenced to date, among them *Cf-2*, *Cf-4*, *Cf-5*, *Cf-9*, *Pto*, *Mi*, *I2*, and *Sw5*. These cloned R genes now provide new tools for tomato breeders to improve the efficiency of breeding strategies, via MAS. Although MAS is still not used routinely for improving disease resistance in many important crops (Michelmore, 2003), it is being used by seed companies for improving simple traits in tomato (Foolad and Sharma, 2005). Furthermore, while the deep knowledge of the tomato genome and the availability of a high-density molecular map for this species (Pillen *et al.,* 1996) should provide further opportunities to accelerate breeding through MAS, the time-consuming and expensive process of developing markers associated with genes of inter-

est and the high cost of genotyping large populations has and will continue to limit the use of MAS in most tomato breeding programmes.

The potential of MAS to speed up the breeding of tomato using molecular markers linked to various resistance genes has been examined in the authors' laboratory. The two main goals of the research were to find the most suitable markers, and to test the feasibility of MAS for pyramiding resistance genes in tomato "elite" lines selected for their good processing qualities.

STRATEGIES FOR GENE TRANSFER AND PYRAMIDING

Six tomato genotypes carrying various resistance genes (Table 3) were crossed with tomato "elite" lines previously selected for yield and quality but lacking resistance traits. Each resistant genotype was crossed initially with each "elite" tomato line and various backcross schemes were then carried out starting from different F_1 hybrids. At each backcross generation the screening of resistant genotypes was performed using molecular markers linked to the resistance genes and DNA extracted from young leaves at seedling stage. Only resistant plants were then transplanted and grown in the greenhouse. At flowering, crosses were made with the recurrent parent to obtain the subsequent generations.

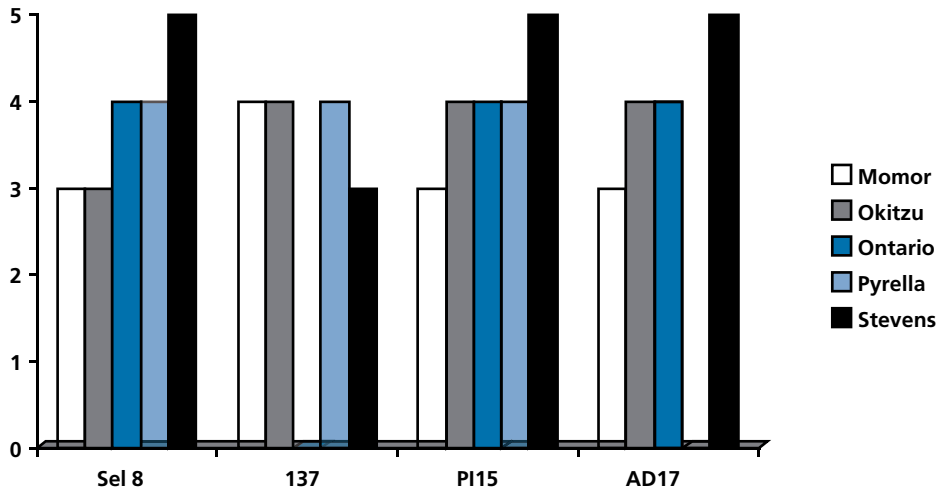
As the efficiency of MAS depends on the availability of polymerase chain reaction (PCR)-based markers highly linked to the resistance gene to be selected, for each resistance gene the most suitable marker system was investigated. For this purpose, three different strategies were undertaken. The first involved searching PCR markers already available in the literature and verifying their usefulness on the genetic material used. The second consisted of designing PCR primers from the sequence of cloned genes reported in the GeneBank database of the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/Genbank), while the third involved designing PCR primers from RFLP markers tightly linked to resistance genes. This last strategy was made possible by the availability of sequences of various mapped tomato RFLPs in the SolGenes database (www.sgn.cornell.edu).

In most cases, the results were obtained using cleaved amplified polymorphic sequences (CAPS; Konieczyn and Ausubel, 1993), which require one PCR reaction followed by restriction digest of the amplified fragment. In three cases (markers linked to genes *Mi*, *Sw5* and *Tm2a*), the primers and enzymes used were those reported in the literature (Williamson *et al*., 1994; Folkertsma *et al*., 1999; Sobir *et al*., 2000). In the case of gene *py-1*, the procedure reported in the literature (Doganlar *et al*., 1998) was

TABLE 3
Tomato genotypes used as resistant parents in the backcross breeding schemes. For each genotype resistant genes are reported

Genotype	Resistance gene	Pathogen
Momor	*Frl, Tm2a, Ve*	*Fusarium oxysporum* f. sp. *radicis-lycopersici,* TMV, *Verticillium dahliae*
Motelle	*I2, Mi, Ve*	*Fusarium oxysporum* f. sp *lycopersici, Meloidogyne* spp., *Verticillium dahliae*
Okitzu	*I2, Tm2a*	*Fusarium oxysporum* f. sp *lycopersici*, TMV
Ontario	*Pto*	*Pseudomonas syringae*
Pyrella	*py-1*	*Pyrenochaeta lycopersici*
Stevens	*Sw5*	TSWV

FIGURE 1
Number of generations reached in backcross schemes carried out between four susceptible recurrent genotypes and five resistant donor genotypes



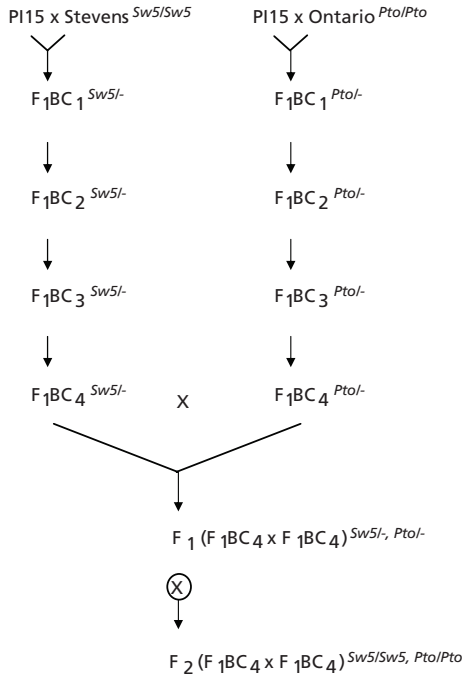
simplified, enabling a faster and cheaper marker system, i.e. a sequence characterized amplified region (SCAR; Kawchuk, Hachey and Lynch, 1998) marker, which only requires one PCR reaction to detect polymorphism between the resistant and the susceptible genotypes, to be set up. (Barone *et al*., 2004).

The second strategy was followed to design primers and enzymes suitable for targeting three resistance genes (*I2, Pto* and *Ve2*). This strategy allowed gene-assisted selection to be achieved through the simple PCR procedure. Finally, the third strategy was applied in the case of one CAPS marker targeting the resistance gene *Frl;* it was derived from one RFLP tomato marker (TG101) linked to the gene (Fazio, Stevens and Scott, 1999).

The markers found were used to select resistant genotypes in backcross breeding schemes, while the process itself allowed three generations to be screened annually. At present, for some cross combinations, the BC_5 generation has been reached, for others the BC_2-BC_3 (Figure 1). Where a BC_5 generation was already available, the breeding programme continued by selfing BC_5 resistant genotypes. In all other cases the backcross programme will continue up to the fifth backcross generation. At the end of each backcross scheme, the resistant BC_5F_3 genotypes, selected through molecular marker analysis, will also be tested directly for resistance by inoculating the pathogen and monitoring signs of disease. This will allow verification that no linkage breakage and loss of resistance gene occurred.

This procedure was already adopted in the case of one backcross scheme aimed at transferring a resistance gene to tomato spotted wilt virus (TSWV) to the susceptible genotype AD17 (Langella *et al*., 2004). The *in vivo* test performed on F_1BC_5, F_2BC_5 and, F_3BC_5 generations confirmed the introgression of the resistance trait and revealed that the resistance gene *Sw5* was

FIGURE 2
Breeding scheme used for the pyramiding of two resistant homozygous genes (*Sw5* and *Pto*) in the susceptible genotype PI15



fixed at the homozygous stage at the F_3BC_5 generation.

Finally, besides the transfer of one resistance gene to each susceptible genotype, a crossing scheme was undertaken to accumulate two or three resistance genes in the same genotype. In this case, the decision was made to stop the backcross scheme at the BC_3 or BC_4 generation as both parental lines were cultivated varieties and hence genetically very similar, and therefore the recovery of the recurrent genome could be satisfactory. F_1BC_4 hybrids carrying the same genetic background in the recurrent parent have been intercrossed, following the breeding scheme shown in Figure 2. At the end of each F_1BC_4 x F_1BC_4 cross and after selecting the genotypes carrying all the resistant alleles at the heterozygous level, one or two selfing generations will be carried out to fix all the resistant genes at the homozygous level.

This strategy has already started in some cases and the first homozygous multiresistant genotypes have been obtained. Also available are two F_2 genotypes out of 52 analysed plants, obtained by intercrossing the F_1BC_4 progeny from PI15 x Stevens with the F_1BC_4 progeny from PI15 x Ontario (Table 4). This F_2 generation exhibited two genotypes carrying both resistant genes *Sw5* and *Pto* at the homozygous level as well as 29 genotypes carrying both genes at the heterozygous level.

The work reported here on transferring resistance genes among tomato genotypes demonstrates the usefulness of MAS for improving traditional breeding strategies. The contribution of molecular markers linked to resistance genes was very efficient in reducing the time and space necessary for selection, enabling both early screening for resistance and reduced numbers of genotypes to be transplanted. The most challenging work was the search for suitable markers, which often required both considerable time and financial resources. Different strategies were used successfully to find the most suitable markers to perform MAS for transferring eight resistance genes into superior tomato genotypes; such strategies could be repeated in tomato for many other genes due to advanced molecular knowledge of the genome of this species.

PERSPECTIVES

The availability of PCR-based markers for many resistance genes allows MAS for biotic resistance in tomato to be applied successfully in any laboratory without the need for highly sophisticated techniques.

TABLE 4
Resistant heterozygous plants obtained in some cross combinations realized for pyramiding of resistance genes

Cross	Pyramided genes	Generation	Analysed plant (number)	Resistant plant (number)
(137 x Momor) F_1BC_3 x (137 x Ontario) F_1BC_3	*Tm2a – Frl* *Ve – Pto*	F_1	50	7
(AD17 x Okitzu) F_1BC_4 x (AD17 x Stevens) F_1BC_4	*Tm2a – Frl* *Sw5*	F_1	24	5
(AD17 x Ontario) F_1BC_4 x (AD17 x Stevens) F_1BC_4	*Pto – Sw5*	F_1	24	6
(PI15 x Stevens) F_1BC_4 x (PI15 x Ontario) F_1BC_4	*Sw5 - Pto*	F_2	52	29

Indeed, once a marker has been set up, its use on large populations for resistance screening is then routine. Technical facilities are today available for screening many samples simultaneously and also costs for equipment are decreasing. In addition, the rapid development of new molecular techniques, combined with the ever-increasing knowledge about the structure and function of resistance genes (Hulbert *et al.*, 2001), will help to identify new molecular markers for MAS, such as single nucleotide polymorphisms (SNPs). Moreover, thanks to the International *Solanaceae* Genome Project (SOL), sequencing of the tomato genome is in progress, and in a few years this will enhance information on resistance genes. This in turn will facilitate the development of molecular markers from transcribed regions of the genome, thereby allowing large-scale gene-assisted selection (GAS) to be achieved.

Over the coming years, selection for pathogen resistance in tomato will be underpinned by research aimed at: mapping other resistance genes for new pathogens; developing PCR-based functional markers (Andersen and Ludderstedt, 2003); designing the most suitable breeding schemes (Peleman and van der Voort, 2003), especially for transferring QTL resistances; large-scale screening through automation; allele-specific diagnostics (Yang *et al.*, 2004); and DNA microarrays (Borevitz *et al.*, 2003). In effect, the combination of new knowledge and new tools will lead to changes in the strategies used for breeding by exploiting the potential of integrating "omics" disciplines with plant physiology and conventional plant breeding, a process that will drive the evolution of MAS into genomics-assisted breeding (Morgante and Salamini, 2003; Varshney, Graner and Sorrells, 2005).

ACKNOWLEDGEMENTS

The authors wish to thank Mr Mark Walters for editing the manuscript and Angela Cozzolino for her technical assistance. This research was partially financed by the Ministry of Agriculture and Forestry of Italy in the framework of project PROMAR (Plant Protection by the Use of Molecular Markers). Contribution no. 119 from the Department of Soil, Plant and Environmental Sciences.

REFERENCES

Ammiraju, J.S.S., Veremis, J.C., Huang, X., Roberts, P.A. & Kaloshian, I. 2003. The heat-stable root-nematode resistance gene Mi-9 from *Lycopersicon peruvianum* is localized on the short arm of chromosome 6. *Theor. Appl. Genet.* 106: 478–484.

Andersen, J.R. & Lubberstedt, T. 2003. Functional markers in plants. *Trends Plant Sci.* 8: 554–560.

Bai, Y., Huang, C-C., van der Hulst, R., Meijer-Dekens, F., Bonnema, G. & Lindhout, P. 2003. QTLs for tomato powdery mildew resistance (*Oidium lycopersici*) in *Lycopersicon parviflorum* G1.1601 co-localize with two qualitative powdery mildew resistance genes. *Mol. Plant - Microbe Interact*. 16: 169–176.

Bai, Y., van der Hulst, R., Huang, C-C., Wei, L., Stam, P. & Lindhout, P. 2004. Mapping Ol-4, a gene conferring resistance to *Oidium neolycopersici* and originating from *Lycopersicon peruvianum* LA2172, requires multi-allelic, single locus markers. *Theor. Appl. Genet.* 109: 1215–1223.

Balint-Kurti, P.J., Dixon, M.S., Jnes, D.A., Norcott, K.A. & Jones, J.D.G. 1994. RFLP linkage analysis of the Cf-4 and Cf-9 genes for resistance to *Cladosporium fulvum* in tomato. *Theor. Appl. Genet.* 88: 691–700.

Ballvora, A., Pierre, M., van den Ackerveken, G., Chornack, S., Rossier, O., Ganal, M., Lahaye, T. & Bonas, U. 2001 Genetic mapping and functional analysis of the tomato *Bs4* locus governing recognition of the *Xanthomonas campestris* pv.*vesicatoria* AvrBs4 protein. *Mol. Plant-Microbe Interact.* 14: 629–638.

Barone, A., Langella, R., Ercolano, M.R., Di Matteo, A. & Monti, L. 2004 Strategies for pathogen resistance gene transfer in tomato with the aid of molecular markers. *In* S.G. Pandalai, ed. *Recent research developments in genetics and breeding*, pp. 211–222. Kerala, India, Research Signpost.

Behare, J., Laterrot, H., Sarfatti, M. & Zamir, D. 1991. RFLP mapping of the *Stemphylium* resistance gene in tomato. *Mol. Plant-Microbe Interact.* 4: 489–492.

Borevitz, J.O., Liang, D., Plouffe, D., Chang, H-S., Zhu, T., Weigel, D., Berry, C.C., Winzeler, E. & Chory, J. 2003. Large-scale identification of single-feaqture polymorphisms in complex genomes. *Genome Res.* 13: 513–523.

Bournival, B.L., Vallejos, C.E. & Scott, J.W. 1990. Genetic analysis of resistances to race 1 and 2 of *Fusarium oxysporum* f. sp. *lycopersici* from the wild tomato *Lycopersicon pennellii*. *Theor. Appl. Genet.* 79: 641–645.

Brouwer, D.J., Jones, E.S. & St Clair, D.A. 2004. QTL analysis of quantitative resistance to *Phytophthora infestans* (late blight) in tomato and comparisons to potato. *Genome* 47: 475–492.

Chagué, V., Mercier, J.C., Guénard, M., de Courcel, A.G.L. & Vedel, F. 1997. Identification of RAPD markers linked to a locus involved in quantitative resistance to TYLCV in tomato by bulked segregant analysis. *Theor. Appl. Genet.* 95: 671–677.

Chunwongse, J., Bunn, T.B., Crossman, C., Jing, J. & Tanksley, S.D. 1994. Chromosomal localization and molecular-marker tagging of the powdery mildew resistance gene (Lv) in tomato. *Theor. Appl. Genet.* 89: 76–79.

Chunwongse, J., Chunwongse, C., Black, L. & Hanson, P. 2002. Molecular mapping of the *Ph-3* gene for late blight resistance in tomato. *J. Hort. Sci. & Biotech.* 77: 281–286.

Danesh, D., Aarons, S., McGill, G.E. & Young, N.D. 1994. Genetic dissection of oligenic resistance to bacterial wilt in tomato. *Mol. Plant-Microbe Interact.* 7: 464–471.

De Giovanni, C., Dell'Orco, P., Bruno, A., Ciccarese, F., Lotti, C. & Ricciardi, L. 2004. Identification of PCR-based markers (RAPD, AFLP) linked to a novel powdery mildew resistance gene (ol-2) in tomato. *Plant Sci.* 166: 41–48.

Diwan, N., Fluhr, R., Eshed, Y., Zamir, D. & Tanksley, S.D. 1999. Mapping of Ve in tomato: a gene conferring resistance to broad-spectrum pathogen, *Verticillium dahliae* race 1. *Theor. Appl. Genet.* 98: 315–319.

Doganlar, S., Dodson, J., Gabor, B., Beck-Bunn, T., Crossman, C. & Tanksley, S.D. 1998 Molecular mapping of the py-1 gene for resistance to corky root rot (*Pyranochaeta lycopersici*) in tomato. *Theor. Appl. Genet*. 97: 784–788.

Fazio, G., Stevens, M.R., & Scott, J.W. 1999. Identification of RAPD markers linked to fusarium crown and root rot resistance (*Frl*) in tomato. *Euphytica* 105: 205–210.

Folkertsma, R.T., Spassova, M.I., Prins, M., Stevens, M.R., Hille, J. & Goldbach, R.W. 1999. Construction of a bacterial artificial chromosome (BAC) library of *Lycopersicon esculentum* cv. Stevens and its application to physically map the *Sw5* locus. *Mol. Breed.* 5: 197–207.

Foolad, M.R. & Sharma, A. 2005. Molecular markers as selection tools in tomato breeding. *Acta Hort.* 695: 225–240.

Foolad, M.R., Zhang, L.P., Khan, A.A., Nino-Liu, D. & Lin, G.Y. 2002. Identification of QTLs for early blight (*Alternaria solani*) resistance in tomato using backcross populations of a *Lycopersicon esculentum* x *L. hirsutum* cross. *Theor. Appl. Genet.* 104: 945–958.

Ganal, M.W., Simon, R., Brommonschenkel, S., Arndt, M., Tanksley, S.D. & Kumar, A. 1995. Genetic mapping of a wide spectrum nematode resistance gene (*Hero*) against *Globodera rostochiensis* in tomato. *Mol. Plant-Microbe Interact*. 8: 886–891.

Gardner, R.G. & Shoemaker, P.B. 1999. "Mountain Supreme" early blight-resistant hybrid tomato and its parents NC EBR-3 and NC EBR-4 tomato breeding lines. *Hort. Sci.* 34: 745–746.

Griffiths, P.D. & Scott, J.W. 2001. Inheritance and linkage of tomato mottle virus resistance genes derived from *Lycopersicon chilense* accession LA 1932. *J. Am. Soc. Hort. Sci.* 126: 462–467.

Grube, R.C., Radwanski, E.R. & Jahn, M. 2000. Comparative genetics of disease resistance within the *Solanaceae Genetics* 155: 873–887.

Hanson, P.M., Bernacchi, D., Green, S., Tanksley, S.D. & Muniy-Appa, V. 2000. Mapping a wild tomato introgression associated with tomato leaf curl virus resistance in a cultivated tomato line. *J. Am. Soc. Hort. Sci.* 125: 15–20.

Huang, C.C., van de Putte, P.M., Haanstra-van der Meer, J.G., Mejer-Dekens, F. & Lindhout, P. 2000. Characterization and mapping of resistance to Oidium lycopersicum in two *Lycopersicon hirsutum* accessions: evidence for close linkage of two Ol-genes on chromosome 6. *Heredity* 85: 511–520.

Hulbert, S.H., Webb, C.A., Smith, S.M. & Sun, Q. 2001. Resistance gene complexes: evolution and utilization. *Ann. Rev. Phytopathol*. 39: 285–312.

Jones, D.A., Dickinson, M.J., Balint-Kurt, P.J., Dixon, M.S. & Jones, J.D.G. 1993. Two complex resistance loci revealed in tomato by classical and RFLP mapping of *Cf-2*, *Cf-4*, *Cf-5*, and *Cf-9* genes for resistance to *Cladosporium fulvum*. *Mol. Plant-Microbe Interact*. 6: 348–357.

Kabelka, E., Franchino, B. & Francio, D.M. 2002. Two loci from Lycopersicon hirsutum LA407 confer resistance to strains of *Clavibacter michiganense* subsp. *michiganensis*. *Phytopath*. 92: 504–510.

Kawchuck, L.M., Hachey, J. & Lynch, D.R. 1998. Development of sequence characterized DNA markers linked to a dominant *Verticillium* wilt resistance gene in tomato. *Genome* 41: 91–95.

Kawchuk, L.M., Hachey, J., Lynch, D.R., Kulcsar, F., van Rooijen, G., Waterer, D.R., Robertson, A., Kokko, E., Byers, R., Howar, R.J., Fischer, R. & Prufer, D. 2001. Tomato *Ve* disease resistance genes encode cell surface-like receptors. *Proc. Nat. Acad. Sci. USA* 98: 6511–6515.

Konieczyn, A. & Ausubel, F.M. 1993. A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* 4: 403–410.

Langella, R., Ercolano, M.R., Monti, L., Frusciante, L. & Barone, A. 2004 Molecular marker assisted transfer of resistance to TSWV in tomato elite lines. *J. Hort. Sci. & Biotech*. 79: 806–810

Laterrot, H. 1996. Breeding strategies for disease resistance in tomatoes with emphasis to the tropics: current status and research challenges, pp.126–132. *Proc.1st Internat. Symp. Trop. Tomato Dis*. Recife, Brazil.

Lefebvre, V. & Chèvre, A.M. 1995. Tools for marking plant disease and pest resistance genes: a review. *Agronomie* 15: 3–19.

Leister, D., Ballvora, A., Salamini, F. & Gebhardt, C. 1996. PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nature Genet*. 14: 421–429.

Levesque, H., Vedel, F., Mathieu, C. & de Courcel, A.G.L. 1990. Identification of a short rDNA spacer sequence highly specific of a tomato line containing Tm-1 gene introgressed from *Lycopersicon hirsutum. Theor. Appl. Genet.* 80: 602–608.

Martin, G.B., Brommonschenkel, S., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D. & Tanksley, S.D. 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262: 1432–1436.

Michelmore, R.W. 2003. The impact zone: genomics and breeding for durable disease resistance. *Curr. Opin. Plant Biol*. 6397–6404.

Moreau, P., Thoquet, P., Olivier J., Laterrot, H. & Grimsley, N.H. 1998. Genetic mapping of Ph-2, a single locus controlling partial resistance to Phytophthora infestans in tomato. *Mol. Plant-Microbe Interact.* 11: 259–268.

Morgante, M. & Salamini, F. 2003. From plant genomics to breeding practice. *Curr. Opin. Biotechnol*. 14: 214–219.

Ori, N., Eshed, Y., Paran, I., Prestino, G., Aviv, D., Tanksley, S.D., Zamir, D. & Fluhr, R. 1997. The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* 9: 521–532.

Parrella, G., Ruffel, S., Moretti, A., Morel, C., Palloix, A. & Caranta, C. 2002. Recessive resistance genes against potyviruses are localized in colinear genomic regions of the tomato (*Lycopersicon* spp.) and pepper (*Capsicum* spp.) genomes. *Theor. Appl. Genet*. 105: 855–861.

Parrella, G., Moretti, A., Gognalons, P., Lesage, M-L., Marchoux, G., Gebre-Selassie, K. & Carata, C. 2004. The Am gene controlling resistance to alfalfa mosaic virus in tomato is located in the cluster of dominant resistance genes on chromosome 6. *Phytopath*. 94: 345–350.

Peleman, J.D. & van der Voort, J.R. 2003. Breeding by design. *Trends Plant Sci.* 8: 330–334.

Pillen, K., Pineda, O., Lewis, C.B. & Tanksley, S.D. 1996. Status of genome mapping tools in the taxon *Solanaceae. In* A. Paterson, ed. *Genome mapping in plants*, pp. 281–308. Austin, TX, USA, Landes.

Rao, A.V. & Agarwal, S. 2000. Role of antioxidant lycopene in cancer and heart disease. *J. Am. Coll. Nutr*. 19: 563–569.

Robert, V.J.M., West, M.A.L., Inai, S., Caines, A., Arntzen, L., Smith, J.K. & St. Clair, D.A. 2001. Marker-assisted introgression of blackmold resistance QTL alleles from wild *Lycopersicon cheesmanii* to cultivated tomato (*L. esculentum*) and evaluation of QTL phenotypic effects. *Mol. Breed.* 8: 217–233.

Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H.S., Lavello, D.T., Dahlbeck, D. & Staskawicz, B.J. 1996. Tomato Prf is a member of the leucine-rich repeat class of

plant disease resistance genes and lies embedded within the tomato Pto kinase gene cluster. *Cell* 86: 123–133.

Sandbrink, J.M., van-Ooijen, J.W., Purimahua, C.C., Vrielink, M., Verkerk, R., Zabel, P. & Lindhout, P. 1995. Localization of genes for bacterial canker resistance in *Lycopersicon peruvianum* using RFLPs. *Theor. Appl. Genet.* 90: 444–450.

Sarfatti, M., Abu-Abied, M., Katan, J. & Zamir, D. 1991. RFLP mapping of I1, a new locus in tomato conferring resistance to *Fusarium oxysporum* f. sp. *lycopersici* race 1. *Theor. Appl. Genet.* 80: 22–26.

Scott, J.W. 2005. Perspectives on tomato disease resistance breeding: past, present, and future. *Acta Hort*. 695: 217–224.

Sobir, Ohmori T., Murata, M. & Motoyoshi, F. 2000. Molecular characterization of the SCAR markers tightly linked to the *Tm-2* locus of the genus *Lycopersicon*. *Theor. Appl. Genet.* 101: 64–69.

Stamova, B.S. & Chetelat, R.T. 2000. Inheritance and genetic mapping of cucumber mosaic virus resistance introgressed from *Lycopersicon chilense* into tomato. *Theor. Appl. Genet.* 101: 527–537.

Stevens, M.R., Lamb, E.M. & Rhoads, D.D. 1995. Mapping the Sw-5 locus for tomato spotted wild virus resistance in tomatoes using RAPD and RFLP analyses. *Theor. Appl. Genet.* 90: 451–456.

Tanksley, S.D. & Costello, W. 1991. The size of *L. pennellii* chromosome 7 segment containing the I-3 gene in tomato breeding lines measured by RFLP probing. *Rep. Tomato Genet. Coop.* 41: 60.

Tanksley, S.D., Young, N.D., Paterson, A.H. & Bonierbale, M.W. 1989. RFLP mapping in plant breeding: new tools for an old science. *BioTech.* 7: 257–264.

Thoquet, P., Olivier, J., Sperisen, C., Rogowsky, P., Laterrot, H. & Grimsley, N.H. 1996. Quantitative trait loci determining resistance to bacterial wilt in tomato cultivar Hawaii 7996. *Mol. Plant-Microbe Interact*. 9: 826–836.

Vakalounakis, D.J., Laterrot, H., Moretti, A., Ligoxigakis, E.K. & Smardas, K. 1997. Linkage between *Fr-1* (*Fusarium oxysporum* f.sp. *radici-lycopersicy* resistance) and *Tm-2* (tobacco mosaic virus resistance-2) loci in tomato (*Lycopersicon esculentum*). *Ann. Appl. Biol*. 130: 319–323.

van der Biezen, E.A., Glagotlkaya, T. & Overduin, B. 1995. Inheritance and genetic mapping of resistance to *Alternaria alternata* f. sp. l*ycopersici* in *Lycopersicon pennellii*. *Mol. Gen. Genet.* 247: 453–461.

van Heusden, A.W., Koornneef, M., Voorrips, R.E., Bruggemann, W., Pet, G., Vrielink-van Ginkel, R., Chen, X. & Lindhout, P. 1999. Three QTLs from *Lycopersicon peruvianum* confer a high level of resistance to *Clavibacter michiganensis* subsp. *michiganensis*. *Theor. Appl. Genet.* 99: 1068–1074.

Varshney, R.K., Graner, A. & Sorrells, M.E. 2005. Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* 10: 621–630.

Williamson, V.M., Ho, J-Y., Wu, F.F., Miller, N. & Kaloshian, I. 1994. A PCR-based marker tightly linked to the nematode resistance, *Mi*, in tomato. *Theor. Appl. Genet.* 87: 757–763.

Yaghoobi, J., Kaloshian, I., Wen, Y. & Willimson, V.M. 1995. Mapping of a new nematode resistance locus in *Lycopersicon* peruvianum. *Theor. Appl. Genet.* 91: 457–464.

Yang, W., Bai, X., Kabelka, E., Eaton, C., Kamoun, S., van der Knapp, E. & Francis, D. 2004. Discovery of single nucleotide polymorphisms in *Lycopersicon esculetum* by computer aided analysis of expressed sequence tags. *Mol. Breed.* 14: 21–34.

Young, N.D. & Tanksley, S.D. 1988. Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the Tm-2 gene in tomato. *Genetics* 120: 579–589.

Yu, Y.G., Wang, J.F., Stall, R.E. & Allejos, C.E. 1995. Genomic localization of tomato genes that control a hypersensitivity reaction to *Xanthomonas campestris* pv *vesicatoria* (Doidge) dye. *Genetics* 141: 675–682.

Zamir, D., Ekstein-Michelson, I., Zakay, U., Navot, N., Zeidan, M., Sarfatti, M., Eshed, Y., Harel, E., Pleban, T., van-Oss, H., Kedar, N., Rabinowitch, H.D. & Czosnek, H. 1994. Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, Ty-1. *Theor. Appl. Genet.* 88: 141–146.

Zhang, L.P., Khan, A., Nino-Liu, D. & Foolad, M.R. 2002. A molecular linkage map of tomato displaying chromosomal locations of resistance gene analogs based on a *Lycopersicon esculentum* x *Lycopersicon hirsutum* cross. *Genome* 45: 133–146.

Zhang, L.P., Lin, G.Y., Nino-Liu, D. & Foolad, M.R. 2003. Mapping QTLs conferring early blight (*Alternaria solani*) resistance in a *Lycopersicon esculentum* x *L. hirsutum* cross by selective genotyping. *Mol. Breed.* 12: 3–19.