Integrating marker-assisted selection in crop breeding – Prospects and challenges

R. Babu[1,*], Sudha K. Nair[1], B. M. Prasanna[2] and H. S. Gupta[1]

[1]Vivekananda Parvatiya Krishi Anusandhan Sansthan (ICAR), Almora, Uttaranchal 263 601, India
[2]Division of Genetics, Indian Agricultural Research Institute, New Delhi 110 012, India

Mapping and tagging of agriculturally important genes have been greatly facilitated by an array of molecular markers in crop plants. Marker-assisted selection (MAS) is gaining considerable importance as it would improve the efficiency of plant breeding through precise transfer of genomic regions of interest (foreground selection) and accelerating the recovery of the recurrent parent genome (background selection). MAS has been more widely employed for simply inherited traits than for polygenic traits, although there are a few success stories in improving quantitative traits through MAS. The success of MAS depends upon several critical factors, including the number of target genes to be transferred, the distance between the flanking markers and the target gene, the number of genotypes selected in each breeding generation, the nature of germplasm and the technical options available at the marker level. With the advent of third-generation marker technologies, such as the single nucleotide polymorphisms, the power and efficiency of genotyping are expected to improve in the coming decades. The present review discusses the basic requirements and the potential applications of MAS in crop plants, the design parameters in a MAS scheme, recent developments in MAS strategies and genotyping techniques, and the significance of integrating MAS into conventional plant breeding programmes.

CONVENTIONAL plant breeding is primarily based on phenotypic selection of superior individuals among segregating progenies resulting from hybridization. Although significant strides have been made in crop improvement through phenotypic selections for agronomically important traits, considerable difficulties are often encountered during this process, primarily due to genotype – environment interactions. Besides, testing procedures may be many times difficult, unreliable or expensive due to the nature of the target traits (e.g. abiotic stresses) or the target environment.

Molecular marker-assisted selection, often simply referred to as marker-assisted selection (MAS) involves selection of plants carrying genomic regions that are involved in the expression of traits of interest through molecular markers. With the development and availability of an array of molecular markers and dense molecular genetic maps in crop plants, MAS has become possible for traits both governed by major genes as well as quantitative trait loci (QTLs). A generalized and simplified key for distinction between qualitative and quantitative traits[1] is presented in Table 1. If the individual gene(s) or QTLs with significant influence on specific target trait(s) can be identified based on their linkage to molecular makers, the efficiency of incorporating the desired traits in elite germplasm could be greatly enhanced. While plenty of information is already available on different kinds of marker systems and gene–marker associations, the practicalities of designing an MAS strategy, increasing their chances of success and efficiency have not received adequate attention till now. A number of molecular investigations have revealed useful gene–marker associations in different crops, but routine implementation in ongoing plant breeding programmes is still in its infancy. It is an opportune time to discuss the practical aspects of designing an MAS strategy and high throughput genotyping techniques that increase their efficiency significantly.

Salient requirements for MAS

In general, the success of a marker-based breeding system depends on three main factors: (i) A genetic map with an adequate number of uniformly-spaced polymorphic markers to accurately locate desired QTLs or major gene(s); (ii) Close linkage between the QTL or a major gene of interest and adjacent markers; (iii) Adequate recombination between the markers and rest of the genome; and (iv) An ability to analyse a larger number of plants in a time- and cost- effective manner.

The success of MAS depends on location of the markers with respect to genes of interest. Three kinds of relationships between the markers and respective genes could be distinguished[2]; (i) The molecular marker is located within the gene of interest, which is the most favourable situation for MAS and in this case, it could be ideally referred to as gene-assisted selection. While this kind of relationship is the most preferred one, it is also difficult to find this kind of markers. For instance, microsatellite or simple sequence repeat (SSR) markers have been

*For correspondence. (e-mail: rbabu_2002@yahoo.com)

designed using the available DNA sequence information for the *opaque2* allele that confers high lysine and tryptophan content in the maize kernel. This has offered an efficient means of tracking the *opaque2* allele in breeding for nutritionally superior maize genotypes, since the marker is located within the gene sequence itself and co-segregates with the target gene. (ii) The marker is in linkage disequilibrium (LD) with the gene of interest throughout the population. LD is the tendency of certain combination of alleles to be inherited together. Population-wide LD can be found when markers and genes of interest are physically close to each other. Selection using these markers can be called as LD-MAS. (iii) The marker is in linkage equilibrium (LE) with the gene of interest throughout the population, which is the most difficult and challenging situation for applying MAS.

However, in most cases, especially for the polygenic traits, the target gene(s) within a QTL have not been characterized at the molecular level. Therefore, genomic regions to be selected using MAS are often chromosome segments carrying QTLs in case of polygenic traits. It is preferable either to have two polymorphic DNA markers flanking the target gene (or a QTL), or a marker within a QTL (if the chromosome segment is more than 20 cM) to eliminate the possibility of genotypes presenting a double recombination between the two flanking markers. Depending on the nature of the genomic region (cloned gene, major or minor QTL) involved in the expression of a target trait and the number of selected QTLs or genomic regions that need to be manipulated, several MAS schemes have been proposed that will be discussed later.

In the context of MAS, DNA-based markers can be effectively utilized for two basic purposes: (i) tracing favourable allele(s) (dominant or recessive) across generations and (ii) identifying the most suitable individual(s) among the segregating progeny, based on allelic composition across a part or the entire genome.

Foreground selection and background selection

Molecular markers are now increasingly being employed to trace the presence of target genes (foreground selection) as well as for accelerating the recovery of the recurrent parent genome (background selection) in backcross programmes. Marker-assisted backcrossing (MAB) improves the efficiency of backcross breeding in three ways: (i) If the phenotype of the desired gene cannot be easily assayed, backcross (BC) progeny possessing a marker allele from the donor parent at a locus near/within the target gene can be selected with a good probability of carrying the gene; (ii) Markers can be used to select BC progeny with least amounts of donor parent germplasm in the genome outside the target region; and (iii) Markers can be used to select rare progeny that are the result of recombination near the target gene, thus minimizing the effects of linkage drag.

Transfer of recessive genes through conventional breeding requires additional selfing generations after every backcross, a procedure that is prohibitively slow for most commercial breeding purposes. Marker-assisted foreground selection was effectively utilized for introgression of disease resistance genes by Melchinger[3], who presented an a priori approach for calculating the minimum number of individuals and family size required in recurrent backcrossing. However, due to lack of allele-specific markers, practical examples of this approach in plant breeding are limited. One successful example is the conversion of normal maize lines into quality protein maize (QPM) through marker-assisted transfer of a recessive mutant allele, *opaque2*, using allele-specific molecular markers. In animal breeding, availability of an array of allele-specific markers has been facilitating applications of this approach on a commercial scale to eliminate disease and stress-susceptibility genes[4].

'Marker-assisted background selection', a term coined by Hospital and Charcosset[5], was initially proposed by Young and Tanksley[6], and experimented by various researchers[7–10]. This strategy has been used extensively in commercial maize breeding programmes, particularly for selection of lines carrying transgenes conferring herbicide tolerance or insect resistance[11]. Several parameters need to be optimized in the background selection programmes. Flanking markers for the target allele are necessary to remove linkage drag. The optimal distance between the target gene and flanking markers governs the selection intensity that can be exerted. The equations given by Hospital *et al.*[7] and Frisch *et al.*[9] are helpful in determining

Table 1. Qualitative vs quantitative traits (modified from Mackill and Junjian[1])

Trait	Segregation	Percentage of phenotypic variation explained	Example	Classification
Qualitative	Discrete	100	Purple leaf colour, blast resistance in rice, anthocyanin pigmentation and opaque endosperm in maize, etc.	Major gene
Semi-quantitative	Discrete	100	Semidwarfism in rice, *sd1* etc.	Major gene
Quantitative	Continuous	> 50	Submergence tolerance gene *sub1* in rice, QTLs for most biochemical traits, etc.	Major gene
Quantitative	Continuous	25–50	Stem rot resistance in rice, etc.	Major QTL
Quantitative	Continuous	< 25	QTL for most agronomic and physiological traits	QTL

the number of BC plants that need to be generated and typed with a special set of flanking markers.

A number of gene–marker associations have been reported in crop plants which could potentially be utilized in MAS strategies (Table 2).

MAS for improvement of qualitative traits: constraints in conventional backcross breeding

When the expression of a target trait is controlled by a single gene or by a gene that accounts for a high proportion of the phenotypic variance of the trait, transfer of that specific gene from the donor to the recipient line can lead to significant improvement of the trait. For the introgression of qualitative traits such as pathotype-specific disease resistance or certain quality characteristics, which are typically governed by single dominant/recessive genes, backcross breeding has been used for a long time[12].

Traditional backcrossing programmes are planned on the assumption that the proportion of the recurrent parent genome is recovered at a rate of $1 - (1/2)^{t+1}$ for each of t generations of backcrossing. Thus, after four backcrosses, we expect to recover $1 - (1/2)^5 = 96.9\%$ of the recurrent parent genome. However, any specific BC progeny will deviate from this expectation due to chance and to linkage between the gene from the donor parent being selected for and nearby genes. A good example of surprising amount of the linkage drag that accompanies backcross breeding programmes was reported by Young and Tanksley[6], who genotyped the chromosome carrying the *Tm2* disease resistance gene in several tomato cultivars that were developed by introgressing the gene from a wild relative, *Lycopersicon peruvianum* via backcross breeding. They found that even cultivars developed after 20 backcrosses contained introgressed segments as large as 4 cM and one cultivar developed after 11 backcrosses still contained the entire chromosome arm carrying the gene from the donor parent. Besides, for the transfer of a single dominant gene, a minimum of six backcross generations would normally be required to recover 99% of the recurrent parent genome. This procedure is too time-consuming, particularly in the context of the competitive nature of modern hybrid breeding programmes, where the turnover times for new lines and hybrids are fast.

Improving qualitative traits using MAS – some case studies

We describe in brief, three of the successful cases of marker-aided backcrossing intended towards the improvement of qualitative traits.

MAS for resistance to soybean cyst nematode: Disease-resistant phenotypes are often simple and oligogenic in nature. Yet, the difficulties in establishing reliable inoculation and scoring methods can challenge even the best plant pathologist or breeder. An example is the resistance to soybean cyst nematode (SCN; *Heterodera glycines*). A widely used phenotypic assay takes five weeks, extensive greenhouse space and 5–10 h of labour for every 100 plant samples processed. The need of the hour for research groups was to develop tightly linked and breeder-friendly markers for the major SCN resistance genes. Due to consistent research efforts, microsatellite or SSR markers have become available for the *rhg1* gene, which is primarily responsible for SCN resistance[13]. Currently, the SSR marker *sat309* located only 1–2 cM away from the *rhg1* forms the basis of most public breeding efforts[14]. This marker alone can be extremely effective in screening breeding populations. Genotypic selections with *sat309* were 99% accurate in predicting lines that were susceptible in subsequent greenhouse assays. Recently, high throughput genotyping using Taqman™ probes for a polymorphism linked to SCN resistance gene, *Rhg4* has been standardized, which is discussed in detail later in this review.

MAS in developing QPM genotypes: Normal maize protein is inherently deficient in two essential amino acids, namely lysine and tryptophan. QPM refers to maize genotypes in which the *opaque2* gene is introgressed along with endosperm modifiers that provide the hard kernel texture. QPM, thereby, offers twice the amount of lysine and tryptophan with kernel vitreousness comparable to the normal maize. While the favourable *opaque2* allele is recessive in nature, the endosperm modifiers that provide hard kernel texture (instead of the characteristic soft texture conferred by the *opaque2* allele in the absence of modifiers), are polygenic. Introgression of *opaque2* along with endosperm modifiers into elite inbreds is complicated because of three major constraints: (i) each conventional BC generation needs to be selfed to identify the *opaque2* recessive gene, and a minimum of six such BC generations are required to recover satisfactory levels of recurrent parent genome; (ii) besides maintaining the homozygous *opaque2* gene, multiple endosperm modifiers must be selected; and (iii) rigorous biochemical tests are required to ensure enhanced lysine and tryptophan levels in the selected materials in each breeding generation, necessitating significant labour, time and resources. Although conventional breeding procedures have been used successfully to convert commercial lines to QPM forms, the procedure is highly cumbersome and time-consuming.

Scientists at CIMMYT (International Center for Wheat and Maize Improvement), Mexico have successfully developed an innovative combination based on SSR markers for *opaque2* allele and phenotypic selection for kernel vitreousness (through the accumulation of endosperm modifiers) for conversion of normal maize lines into QPM. Three SSR markers, *umc1066*, *phi057* and *phi112*,

Table 2. Selected examples of gene–marker associations for important traits in major crops

Crop	Trait	Gene/QTL	Linked marker(s)	Reference
Rice	Blast resistance	*Pi-1*	RZ 536 and r 10	11
		Pi-2	RG 64	18
		Pi-4	RZ 788	56
		Pi-9	RG 16, pB A 14 and pB 8	57
		Pi-44	CD 0520	58
	Bacterial blight resistance	*Xa-1*	XNpb 235	59
		Xa-3	XNpb 181	60
		Xa-4	XNpb 181	60
		xa-5	RZ 390/RG 550/RG 207	60
		xa-13	RG 136	61
		Xa-21	RG 103, pTA 248	62
		Xa-22(t)	R 543/RZ 536	63
	Rice tungro virus resistance	*RTSV*	RZ 262	64
	Gall midge resistance	*Gm 2*	RG 329	65
		Gm 4 (t)	R 1813	66
		Gm 7	SA 598 (AFLP)	67
	Brown plant hopper resistance	*Bph 1*	XNbp 248	68
		Bph (t)	RZ 404/UCH 170	69
	Green leaf hopper resistance	*Grh 1*	R 566	70
	Submergence tolerance	*Sub 1*	RZ 698	71
	Salt tolerance	*OSA 3*	RG 457/ Y 12817R	72
	Wide compatibility	*S 5*	RG 213	73
	Temperature sensitive male sterility	*TGMS*	RM 257 and TS 200	74
	Grain aroma	*Fg r*	RG 28	41
	Amylose content	*Wx*	wx	75
	Photoperiod sensitivity	*Se 1*	RG 64	76
	Semi-dwarf stature	*Sdg*	R 2182	77
	Shattering resistance	*Sh 4*	R 250	78
		Sht	R 1427/C 107	79
Maize	Northern corn blight resistance	*Htn-1*	umc 117	80
	Cytoplasmic male sterility	T, C and S cytoplasms	STS markers derived from mt DNA sequences, viz. M 112582, S 81074 and AF 008647	81
	Enhanced lysine and tryptophan (QPM)	*opaque-2*	umc 1066, phi 112 and phi 057	82
	Days to pollenshed	QTL	Taqman probes	46
Wheat	Cyst nematode resistance	*Ccn-D1*	Cs E 20-2	83
	Leaf rust resistance	*Lr 24*	6 RFLP markers	84
		Lr 35	SCAR marker developed from ISSR	84
		Lr 9	3 RAPD markers and 1 RFLP marker	85
		Lr 47	RAPD marker	86
	Powdery mildew resistance	*Pm-1 and Pm-2*	RFLP marker	87
		MIRE	SSR markers *gwm* 427 and *gwm* 617	88
	Hessian fly resistance	*H 6*	Op AF 08, Op B01 and UBC 511 (STS)	89
	Root lesion nematode resistance	*Rlnn1*	Xedo 347-7A	90
	Earliness per se	*Eps-A^{m}1*	Xwg 241 and Xc do 393	91
	Loose smut resistance	*T 10*	SCAR marker	92
	Vernalization requirement	*Vrn-B1*	$(TG)_3$ and dCAPS primers	93
	Coleoptile colour	*Rc-A1*, *Rc-D1* and *Rc-B1*	Xgwm 913, Xgwm 111 and Xgwm 1184	94
	Flour colour	Major QTL	STS marker	95
Sorghum	Head smut resistance	*shs*	RFLP probes, viz. pFBT, xS 560, xS 1294 and RAPD marker OPG5	96
	Fertility restoration	*rf4*	LW 7, LW 8 and LW 9 (STS – CAPS markers)	97
Soybean	Cyst nematode resistance	*rhg 1*	Sat 309 (SSR marker)	14
	Soybean mosaic virus resistance	*Rsv*	RFLP and microsatellite markers	98
	Linolenic acid content	*Fan*	pB 194-1 and pB 124	99
	Super nodulation ability	*nts*	pA – 132	100
Chickpea	Double-podding	*s*	TA 80 (STMS marker)	101
Pea	Nodulation ability	*Sym 9*	A 5/14, A 5/16 (AFLP)	102
		Sym 10	'chs 2' (RFLP)	
	Powdery mildew resistance	*er*	p 236 (RFLP) and SCAR markers developed from OPD – 10_{650}, OPE 16, OPL 06 and OPO 18	103, 104
	Fusarium wilt resistance	*Fw*	RFLP marker	105
Sunflower	Downy mildew resistance	*P16*	HaP1, HaP2 and HaP3	106
Tomato	*Meloidogyne incognita* resistance	*Mi*	RAPD marker	107
		Mi 3	RAPD and RFLP markers	108
	Yellow leaf curl virus resistance	*Ty 1*	RFLP marker	109
	Black mold resistance	QTL	RFLP markers	110
Potato	*Globodera rostochinensis* resistance	*H 1*	RFLP marker	111
	Resistance to potato virus X	*Nb*	AFLP markers GM 339 and GM 637	112
	Late blight resistance	*R 1* and *R 3*	RFLP marker	113

which are internal repetitive sequences within the *opaque2* gene are being utilized in PCR-based assays to select individuals carrying the *opaque2* gene in successive backcross generations, and hence, the time required is reduced to one-half. In addition, marker-aided background selection can help recover the same level of recurrent parent genome in three generations as would be achieved by six generations of conventional selections. Also, errors due to recombination frequency between the target gene and the flanking marker can be eliminated, since these SSR markers are located within the target gene itself. Apart from this, routine biochemical tests to ensure high lysine and tryptophan levels need not be carried out at each generation since the presence of *opaque2* is confirmed. Thus, MAS based on SSR markers for conversion of normal maize lines into QPM is simple, rapid, accurate, efficient, cost-effective and complementary to existing breeding protocols[15].

Marker-aided pyramiding of rice genes for bacterial blight and blast resistance: The most effective approach to combat bacterial blight caused by *Xanthomonas oryzae* pv. *oryzae* is the use of resistant varieties. So far, 25 resistance genes have been identified[16] and some of them have been incorporated into modern rice varieties through conventional breeding. Transfer of *Xa4* gene resulted in the development of many bacterial blight-resistant rice varieties, which have played a vital role in sustaining the rice yields in many countries. However, large-scale and long-term cultivation of varieties carrying *Xa4* resulted in a significant shift of the dominant bacterial blight pathogens. Rice varieties with single *Xa4* gene have become susceptible to the bacterial blight disease in many areas in Indonesia, India, China and Philippines. One way to delay such a breakdown in bacterial blight resistance is to pyramid multiple resistance genes into rice varieties. This can be difficult or nearly impossible with conventional breeding approach due to epistasis of genes, particularly when a breeding line already carries a gene, e.g. *Xa21* that shows resistance to most of the bacterial blight races at the time the varieties are being developed. With the conventional approach, a breeding line with *Xa21* only cannot be distinguished from a breeding line with *Xa21* plus some other genes. International Rice Research Institute (IRRI), Philippines and Punjab Agricultural University, Ludhiana have successfully employed MAS to pyramid genes for bacterial blight resistance. All possible combinations of the four resistance genes, viz. *Xa4*, x*a5*, x*a13* and *Xa21* were pyramided using STS (sequence tagged site) markers. The pyramided lines show wider spectrum or higher level of resistance to the bacterial pathogen. Recently, the Government of Indonesia released two new rice varieties, 'Angke' and 'Conde' through MAS, which are resistant to bacterial leaf blight infection[17].

Rice blast caused by the fungal pathogen, *Magnaporthe grisea* is another important disease throughout the world. The dynamic evolution of the virulence in the blast fungus coupled with the epistatic action of different resistance genes in the host plant, makes breeding for blast resistance a constant challenge. Molecular breeding approach is being widely employed for the improvement of blast resistance in many high-yielding commercial rice cultivars. Three genes, namely *Pi1*, *Piz5* and *Pita* have been pyramided in a susceptible rice variety, CO39 using RFLP and PCR-based markers for durable blast resistance[18].

MAS for improvement of quantitative traits

Most of the traits of agronomic importance are complex and regulated by several genes. Unlike the case of simply inherited traits that are controlled by one or a few major genes, improvement of polygenic traits through MAS is a complex endeavour. The difficulty in manipulating quantitative traits is related to their genetic complexity, mainly the number of genes involved in their expression and interactions among genes (epistasis). Because several genes are involved in expression of a quantitative trait, these genes, in general, have smaller individual effects on the phenotype, and the effect of the individual genes is not easily identifiable. This warrants repetitions of field tests to characterize accurately the effects of QTLs and to evaluate their stability across environments. The evaluation of QTL by environment interaction (Q × E) continues to be a major constraint on the efficiency of MAS[19]. Besides, epistatic interactions among different regions of a genome can induce a skewed evaluation of QTL effects. Also, if the genomic regions involved in the interactions are not incorporated in the selection scheme, they can potentially bias the selection process.

Despite the proliferation of QTL mapping experiments in recent years, a number of constraints have imposed severe limitations on efficient utilization of QTL mapping information in plant breeding through MAS. Salient among these constraints are: (i) identification of a limited number of major 'players' (QTLs) controlling specific traits; (ii) the notion that QTL identification is required whenever additional germplasm is used; (iii) inadequacies/experimental deficiencies in QTL analysis leading to either overestimation or underestimation of the number and effects of QTLs; (iv) lack of universally valid QTL-marker associations applicable over different sets of breeding materials; (v) strong QTL – environment interaction; and (vi) difficulty in precisely evaluating epistatic effects.

Increasing the efficiency of MAS for quantitative traits calls for improved field experimentations/designs, robust mathematical models and sound statistical approaches. For instance, with composite interval mapping (CIM), field data from different environments can be integrated into a joint analysis to evaluate the Q × E interactions, thus, enabling identification of stable QTLs across environments[20]. Besides, with a detailed linkage map, CIM allows a precise

identification of the QTL in the genome and better identification of linked QTL (in coupling phase) from the same parental line. DNA-based markers also offer the unique advantage of identification and utilization of favourable QTL from even a phenotypically inferior parental line (in repulsion phase).

An especially powerful approach for enhancing the QTL mapping in tandem with MAS was proposed by Tanksley and Nelson[21]. The advanced backcross-QTL (AB-QTL) analysis involves crosses between elite germplasm and unadapted genotype/wild relatives with favourable genes/QTLs, followed by two generations of backcrossing for developing several hundred sibling lines. These lines, each retaining different genomic segments of the wild relative/unadapted genotype, are then genotyped using DNA markers. The lines become, in effect, a set of near-isogenic lines that individually dissect the effects of potential QTL in the background of the elite parent. At the same time, the BC lines also provide relatively fixed material for the critical step of replicated phenotypic evaluations. The essence of this strategy – uncovering and accessing the desirable alleles from wild relatives or unadapted genotype – demonstrates that QTL mapping can go hand in-hand with MAS rather than as sequential steps.

QTL-marker associations however, need to be unravelled through more intensive research efforts. It would be critical to find out whether extremely tight linkages between marker loci and QTL may lead to highly conserved allele associations. If so, observing marker allele frequency changes in long-term selection experiments or determining markers that explain significant portions of the combining ability variance in diallel or factorial crosses, might reveal universally applicable markers[22]. Another approach would be to conduct QTL analyses in genetically broad-based panmictic populations by means of highly saturated integrated genetic marker maps. For a maximal efficiency of MAS, direct QTL-allele-specific markers (such as STS markers derived from cloned QTL alleles) are needed[23]. Although success in terms of cloning of QTL alleles is highly limited, map-based cloning and candidate gene approaches would increasingly facilitate isolation and characterization of agronomically important QTLs[24], particularly due to the rapid progress being made in genome sequencing of various plants. There have been few reports of successful map-based positional cloning of QTLs till recently. QTLs for glucose and fructose content in tomato (*Brix 9-2-5*)[25], fruit weight in tomato (*fw 2.2*)[26], heading date in rice (*Hd1*)[27] and (*Hd6*)[28], etc. have been cloned. Cloning the genetic determinants of QTLs is expected to bridge the missing link in our understanding of the relationship between genotype and phenotype.

Quantitative trait improvement through MAS – some case studies

MAS for improving heterotic performance in maize: The maize lines B73 and Mo17 are not highly productive, but widely used across the world, and represent the two most important 'heterotic groups' in the US maize breeding programmes. Based on a mapping population derived from the B73 × Mo17 cross, QTLs contributing to heterosis for grain yield were mapped on nine of the ten maize chromosomes. Further, mapping studies suggested that two elite inbred lines, Tx303 and Oh43 contained genetic factors that might improve the heterotic response of the B73 × Mo17 single cross hybrid. Six chromosomal segments each in Tx303 and Oh43 were transferred (using MAS) through three backcross generations into the target lines, B73 and Mo17. Two selfing generations followed the third backcross. Based on the initial evaluations of the testcross hybrids, the better-performing modified lines were selected for intercrossing and were designated as 'enhanced' lines. The single cross hybrids derived by crossing the enhanced B73 × enhanced Mo17 exceeded the hybrid 'checks'[29,30] by 12 to 15%. This study was one of the earliest demonstrations that marker-facilitated backcrossing can be successfully employed to improve complex traits such as grain yield in maize.

Germplasm enhancement in tomato using AB-QTL strategy: The effectiveness of AB-QTL strategy was demonstrated in tomato by Tanksley and coworkers through a series of studies[31–34]. Genetic enhancement, through AB-QTL strategy, for various traits of agronomic importance, including fruit quality and black mould resistance in tomato, were accomplished using wild relatives such as *Lycopersicon pimpinellifolium*, *L. peruvianum*, *L. hirsutum* and *L. cheesmanii*.

MAS for drought stress tolerance in maize: CIMMYT researchers have made considerable efforts during the past three decades to improve pre- and post-flowering drought tolerance in maize. Although significant progress has been achieved for improving drought tolerance in CIMMYT maize germplasm through conventional breeding, the approach is slow and time-consuming. Use of molecular markers and QTL information based on carefully managed replicated tests showed the potential to alleviate the problems associated with inconsistent and unpredictable onset of moisture stress or the confounding effect of other stresses such as heat. The complex trait of drought tolerance was first broken down into simpler components, such as anthesis-silking interval, that are closely associated with drought tolerance. CIMMYT then conducted a series of experiments on QTL analysis and MAS for transfer of drought tolerance to tropical maize, and obtained encouraging results. An integrated strategy of QTL-mapping, MAS and functional genomics is now being explored to further provide genomic information and tools to effectively complement conventional selection for improving drought stress tolerance in maize[35].

Designing a BC–MAS scheme: some important considerations

In a BC scheme, the strategy is to transfer a specific favourable allele at a target locus from a donor line to a recipient line. The use of DNA markers, which permit the genetic dissection of the progeny at each generation, increases the speed of selection process[36]. In general, the BC scheme aims either at complete or partial line conversion. The objective of a complete line conversion is to develop a line that will have exactly the same genetic composition as the recipient line, except at target loci where the presence of homozygous alleles from a donor line is desired. Partial line conversion means that the conversion is complete when a limited proportion of the donor genome in an individual is found scattered over the genome in addition to the desirable homozygous alleles at the target gene.

Factors considered critical in a BC–MAS scheme include the number of target genes, the distance between the flanking markers and the target gene (2–20 cM), and the number of genotypes selected in each BC generation. Depending upon the objectives, the experimental design for line conversion through BC–MAS needs to be modulated based on the available resources, nature of the germplasm (e.g. agronomic quality and number of lines to be converted) and technical options available at the marker level. Once the number of target genes to be introgressed has been defined, the next step would be to determine the population size that needs to be screened at each generation, giving a target-selectable population size of 50 to 100 genotypes. Once this is defined, one should determine the desirable recombination frequency between the flanking markers and target gene and the number of genotypes selected at each generation based on the objective and the constraints of the experiment. The number of BC generations required to achieve the introgression can be predicted based on simulations[37]. When resources are limited or introgression from a donor line into a large number of recipient lines is desired, strategies based on BC–MAS at one target locus solely at one advanced BC generation should be considered. Selection in later generations is more effective because the ratio of the standard deviation to the mean of the donor genome contribution increases as the backcrossing proceeds. The backcross procedure can be terminated after four, instead of six, backcross generations even with small population sizes and limited number of marker data points (MDP). Thus, the marker technology can be advantageous even when the resources in a breeding programme are limited. MAS has the potential to reach in BC_3 generation, the same level of recurrent parent genome (RPG) as reached in BC_7, without the use of markers[9]. However, large numbers of MDP and more efficient marker systems are required to unlock this potential.

In the above scheme, it is critical to identify the most convenient set of markers for the allele(s) of interest, because the screening of the whole population has to be conducted at least once at the beginning of each BC generation. With the population size running into hundreds or thousands, such screening can be laborious and expensive. However, this can be optimized using an appropriate combination of DNA markers. With the recent advances in molecular technology, particularly single nucleotide polymorphisms (SNPs)[38], a substantial improvement in the capacity to efficiently screen larger populations can be achieved.

Efficiency of MAS

Computer simulations have provided powerful tools for analysing the design and efficiency of MAS programmes[8]. Three different selection strategies in a marker-assisted background selection programme, namely two-step, three-step and four-step selections, were compared by computer simulation in terms of quicker recovery of a large proportion of the RPG[9]. The simulations were based on maize genetic map ($n = 10$) with markers spaced about 20 cM apart and with the assumption that the target locus could be scored directly either through phenotype or a marker completely linked to the target gene. Major conclusions from this simulation experiment are as follows:

(i) The four-stage sampling strategy which includes (a) selecting individuals carrying the target allele; (b) selecting individuals homozygous for recurrent parent genotype at loci flanking the target locus; (c) selecting the individuals homozygous for recurrent parent genotype at the remaining loci on the same chromosome as the target allele; and (d) selecting one individual that is homozygous for the recurrent parent genotype at most loci (across whole genome) among those that remain, is the most efficient procedure in general.

(ii) With the four-stage sampling strategy and reasonable population size (50–100), one can expect to find BC_3 progeny with at least 96% RPG with 90% probability. It would take six generations of traditional backcrossing to reach this stage, besides the risk of a larger probability of linkage drag around the target gene.

(iii) Increasing the number of markers genotyped at each generation had little effect. Once the threshold of one marker per 20 cM is reached, additional markers (except perhaps around the target locus) would not be required. The frequency of recombination, and not the number of markers, is a more important limiting factor in reducing linkage drag, which suggests that sampling larger populations with fewer markers makes sense than the reverse.

Recent advances in MAS strategies and genotyping techniques

Single large-scale MAS

In this strategy[39], selection of parental lines is first carried out among outstanding elite material for the trait to

be improved with the best allelic complementarity. By crossing each selected parental line with a tester (elite line lacking the target trait), segregating populations are developed. Genomic regions of interest for each parental line are identified by combining favourable alleles in the segregating populations (e.g. F_3 families and RILs). MAS, based on reliable PCR-based markers to fix favourable alleles at target genomic regions, is conducted only once on large segregating populations derived from crosses between the elite lines. The strategy is suggested to offer three distinct advantages: (i) favourable alleles (selected to improve a specific trait) are derived from two or more sources of elite parental lines in a complementary scheme, disregarding the 'recipient/donor' line concept; (ii) plants with fixed favourable alleles at specific genomic regions are selected at an early generation of recombination and no pressure of selection is applied outside the targetted regions. This assures good allelic variability in the rest of the genome for future line development under various conditions and environments; and (iii) the approach is particularly relevant for pyramiding favourable alleles at cloned genes or major QTLs in new germplasm.

Pedigree MAS

This approach[40] is especially relevant for crops such as wheat, where pedigrees of elite germplasm are known. Fingerprinting elite wheat materials must be conducted in a set of lines actively used in the breeding programme, and in elite materials to be subsequently released. The data may be combined with the phenotypic data collected during different selection cycles to identify favourable alleles for trait(s) of interest. For example, if an elite line contains alleles for yield performance in a target environment, their frequency should be higher than the expected random frequency in offspring derived from this elite parental line. This shift in allelic frequency reflects phenotypic selection by breeders and may be identified by comparing fingerprinting data of both parents and their offspring. Once the favourable alleles are identified, DNA markers closely linked to the target genomic regions can be used to accelerate fixation of favourable alleles in the next selection step: a new set of elite materials (offspring 1) to derive the next set of elite lines (offspring 2). This MAS strategy was suggested to be most efficient when conducted on F_2 or F_3 segregating populations[41].

Breeding by design™

'Breeding by design' is a novel concept that aims to control all allelic variations for all genes of agronomic importance. The authors of this concept[42] propose that by understanding the genetic basis of all agronomically important characters and the allelic variation at those loci, the breeder would be able to design superior genotypes *in silico*, which demonstrates that DNA markers not only improve selection processes but can aid in creating novel genotypes bearing new characteristics of agronomic importance. The strategy involves three well-defined steps: (i) mapping loci involved in all agronomically relevant traits, preferably through introgression line (IL) libraries that are extremely powerful not only in mapping, but also in reducing the complexity of polygenic traits by separating them into a set of monogenic loci. An IL library consists of a series of lines harbouring a single homozygous donor segment introgressed into a uniform cultivated background[43]. (ii) Assessment of the allelic variation at those loci through intensive chromosome haplotyping and extensive phenotyping of all agronomic traits. Haplotypes refer to a combination of linked marker alleles that occur in a locus in a set of accessions. (iii) Designing superior genotypes comprising a combination of favourable alleles at all loci through accurate selection of recombination events using flanking markers to collate different favourable alleles next to each other. 'Breeding by design' involves an integrative, complementary application of technological tools and materials currently available to develop superior varieties. However, success of this approach would essentially depend upon availability of extremely saturated marker maps and precision of phenotyping.

Non-gel based separation of PCR-amplified DNA products

The efficiency of MAS in larger breeding populations could be improved if the laboratory throughput is high and costs are low. Currently, the most common way to evaluate the outcome of PCR amplification is to analyse the amplified products using gel electrophoresis once thermal cycling is complete. This process takes extra time and effort and also increases the chance of laboratory equipment or reagents being contaminated. Consequently, attempts have been made to develop means to detect the accumulation of PCR products either during thermal cycling or following thermal cycling without having to remove the samples for processing. With the advent of third-generation markers, such as SNPs, dispensing with electrophoresis as a core technology to detect polymorphism has become the primary focus. All major SNP detection techniques rely on an initial PCR amplification of the target DNA segment, while they differ in the ways in which the alternate amplicon sequences are discriminated from one another. The SNP is targetted either internally within the amplicon (Taqman™, molecular beacon, etc.) or/at immediately downstream of the 3′ end of one of the amplification primers, using allele specific oligonucleotides. The probes employed in both molecular beacon and Taqman™ technologies are based on the principle of fluorescence resonance energy transfer[44,45].

Recently, non-electrophoresis-based PCR assays for the allelic discrimination at two linked loci flanking an important QTL controlling days to pollen shed in maize were described[46]. The assays are based on the fluorogenic 5′-nuclease procedure (Taqman), which allows for direct detection of the PCR product by the release of a fluorescent reporter dye as a result of DNA amplification (Figure 1). The main steps in the reaction sequence are polymerization, strand displacement and cleavage. Two dyes, a fluorescent reporter (*R*) and a quencher (*Q*) are attached to the fluorogenic probe. When both dyes are attached to the probe, reporter dye emission is quenched. During each extension cycle, the DNA polymerase cleaves the reporter dye from the probe. Once separated from the quencher, the reporter dye fluoresces. The assays were developed after sequencing the alleles at both the loci, by designing suitable primers and probes based on SNPs or insertion/deletion polymorphisms. The Taqman procedure allowed for a fast and highly reproducible analysis directly in the PCR vials. The complete automation of the process makes this technique highly valuable for the large-scale screening required for MAS studies and map-based cloning projects. Similarly, high throughput genotyping using Taqman probes for a polymorphism linked to SCN resistance gene, *Rhg4* has been standardized[47]. This method can be used for scoring polymorphism in an RIL population and for scoring parental lines in breeding programmes. The Taqman method of determining genotype was accurate in 90% of scores in the RIL population compared to 95% accuracy with electrophoresis. Therefore, this method can be applied to automated large-scale genotyping for soybean breeding programmes.

The molecular beacons have been demonstrated to be useful in discriminating between transgenic and non-transgenic cereals. The principle of molecular beacon assay is explained in Figure 2. The hairpin stem formed by the complementary arm sequences of the beacon cannot coexist with the double helix that is formed when the probe hybridizes to its target. Consequently, the molecular beacon undergoes a conformational change that forces the arm sequences apart and causes the fluorophore to move away from the quencher, thereby allowing fluorescence. Oligonucleotide ligation assay[48–50] and dynamic allele-specific hybridization[51] have been developed, which can potentially increase the power of DNA genotyping.

Another interesting and innovative assay is invader SNP assay, which is PCR- and gel-free. Invasive cleavage is a novel non-PCR technique relying on the action of a special endonuclease to remove overlapping oligonucleotides during polymerase-induced extension. These assays are performed on total genomic DNA using universal invader and fluorescent probes and an allele-specific signal probe. The invader assay employs endonucleases (derived from *Archaea* species) from the FEN family (cleavases) to cleave structure-specific rather than sequence-specific sites. Allele-specific oligonucleotide probes (ASOP) are designed to produce the structure necessary for cleavage of the signal probe. If ASOP matches the sequence of the allele, then the required flap structure is produced and the signal probe is cleaved (Figure 3). If the ASOP mismatches the sequence of the allele, then the required flap structure is not produced and signal probe is not cleaved[52]. The simple semi-automated data interpretation of the assay allows the screening of 25,000 to 50,000 genotypes per operator per week, which is tenfold greater than gel-

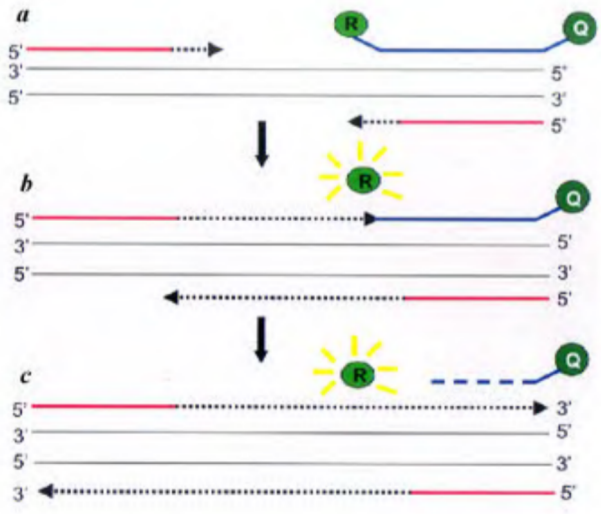

Figure 1. Schematic diagram of Taqman assay. ***a***, Polymerization; ***b***, Strand displacement and cleavage; ***c***, Polymerization completed. *R*, Reporter; *Q*, Quencher.

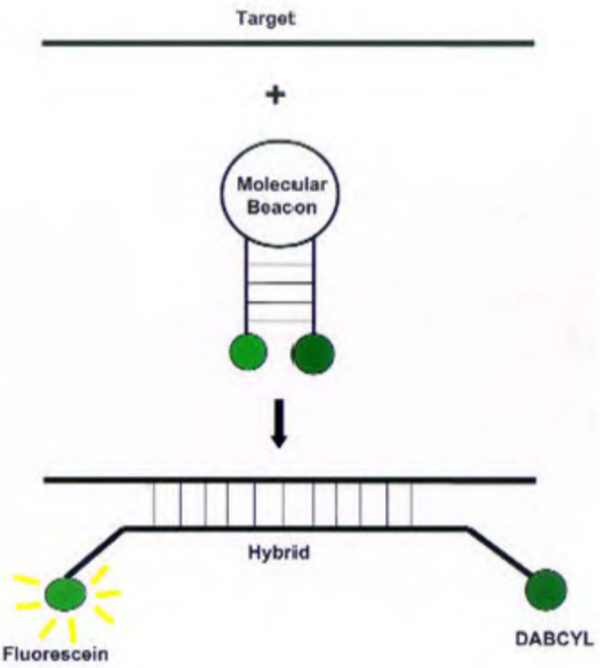

Figure 2. Molecular beacon assay.

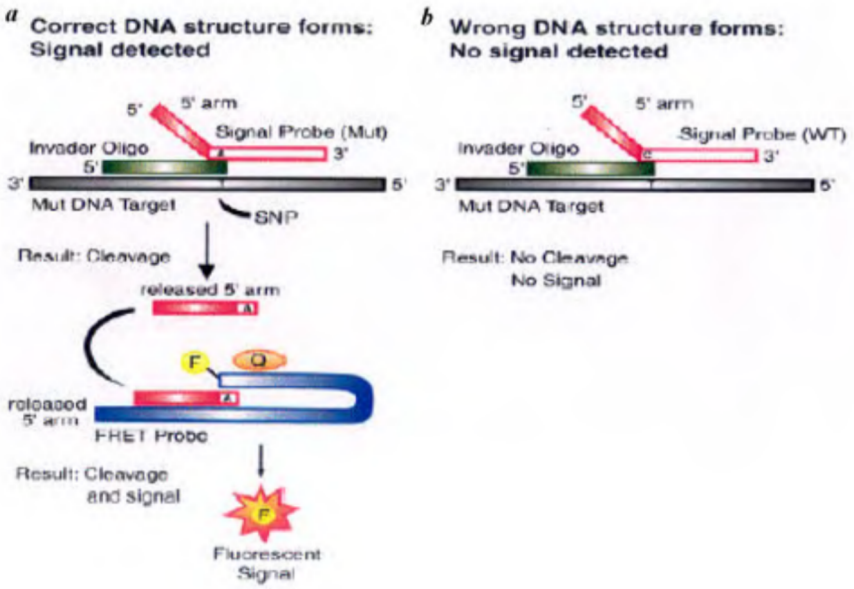

Figure 3. Principle of invader SNP assay. (Source: http://cropandsoil.oregonstate.edu)

based SNP and microsatellite typing. The SNP typing in this assay is found to be highly accurate when compared to standard PCR-based methods. The assay requires no special instruments other than a microplate reader and proves to be a highly sensitive method for detecting SNPs directly from the genomic DNA with high genotyping throughput.

MAS in plant breeding: need for cautious optimism

Integrating molecular marker technologies such as MAS into breeding strategies could become increasingly important in the coming years, to realize genetic gains with greater speed and precision. Although MAS is currently used more widely for simply inherited traits than for polygenic traits, with the development and access to reliable PCR-based markers such as SSRs and SNPs, in several crop plants, efficiency of genotyping large populations or breeding materials has significantly increased. Marker-assisted foreground selection is already gaining rapid momentum as allele-specific markers are now becoming available, through crop genomics research, on a number of agronomically important traits.

The promise of MAS for improving polygenic traits in a quick time-frame and in a cost-effective manner is still elusive. There is a wider appreciation that simply demonstrating that a complex trait can be dissected into QTLs and mapped to approximate genomic locations using DNA markers would not serve the ultimate goal of trait improvement. As suggested by Young[53], research on quantitative traits need to utilize better scoring methods, larger population sizes, multiple replications and environments, appropriate quantitative genetic analysis, various genetic backgrounds and, whenever possible, independent verification through advanced generations or parallel populations. In facing the challenge of improving several lines for quantitative traits, MAS strategies should probably concentrate on using DNA markers in one key selection step to maximize their impact. They could be used at the very beginning of the scheme as predictive tools to reduce the number of crosses, at an early stage of recombination to fix target genomic regions, or at an advanced stage of germplasm development as a diagnostic tool, when the allelic value has already been identified[39].

The field of 'functional genomics' is making rapid strides, aiming towards assessing the gene functions through genome-wide experimental approaches[54]. Innovative tools such as DNA chips, microarrays, expressed sequence tags, etc. help in the quantitative assessment of RNA levels in the sample, which in turn might benefit the breeder in selecting best lines based on RNA expression profiles[55] as much as marker genotypes. These innovative approaches are not only expected to identify more genes involved in regulating different pathways in response to stress conditions, but also mark a significant paradigm shift in the field of molecular breeding. In future, quantitative genetics will look towards genomics for information to develop more biologically meaningful models, while genomics will look towards quantitative genetics to develop and validate hypotheses involving complex gene interactions. Bioinformatics will play an important role in facilitating this crossover. Thus, integrating genomics and bioinformatics into the field of molecular breeding is expected to bring in more fundamental revolutions in plant

breeding. It would be crucial to recognize the complementarities between molecular technologies and conventional breeding and harness this vital synergy into developing comprehensive research strategies aimed towards more efficient crop improvement in future.

1. Mackill, D. J. and Junjian, N., Molecular mapping and marker assisted selection for major gene traits in rice. In *Rice Genetics, IV*, Science Publishers Inc., IRRI, Philippines, 2001, pp. 137–151.
2. www.fao.org\EMAS.
3. Melchinger, A. E., Use of molecular markers in breeding for oligogenic disease resistance. *Plant Breed.*, 1990, **104**, 1–19.
4. Plastow, G. S., Needs for Biotechnology in Animal Production. In Proceedings of the 7th Agrogene Seminar, Paris, France, 1999.
5. Hospital, F. and Charcosset, A., Marker assisted introgression of quantitative trait loci. *Genetics*, 1997, **147**, 1469–1485.
6. Young, N. D. and Tanksley, S. D., RFLP analysis of the size of chromosomal segments retained around Tm-2 locus of tomato during back cross breeding. *Theor. Appl. Genet.*, 1989, **77**, 353–359.
7. Hospital, F. *et al.*, Using markers in gene introgression breeding programs. *Genetics*, 1992, **132**, 1199–1210.
8. Visscher, P. M. *et al.*, Marker assisted introgression in back cross breeding programs. *Genetics*, 1996, **144**, 1923–1932.
9. Frisch, M. *et al.*, Comparison of selection strategies for marker assisted back crossing of a gene. *Crop Sci.*, 1999, **39**, 1295–1301.
10. Frisch, M. *et al.*, Minimum sample size and optimum positioning of flanking markers in marker assisted back crossing for transfer of a target gene. *Crop Sci.*, 1999, **39**, 967–975.
11. Yu, Z. H. *et al.*, Molecular mapping of genes for resistance to rice blast (*Pyricularia grisea Sacc*). *Theor. Appl. Genet.*, 1996, **93**, 859–863.
12. Allard, R. W., *Theor. Principles of Plant Breeding*, Wiley, New York, 1960.
13. Mudje *et al.*, Two microsatellite markers that flank the major soybean cyst nematode resistance locus. *Crop Sci.*, 1997, **37**, 1611–1615.
14. Cregan *et al.*, Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by *rhg1* locus. *Theor. Appl. Genet.*, 2000, **99**, 811–818.
15. Dreher, K. *et al.*, Is marker assisted selection cost-effective compared to conventional plant breeding methods? The case of Quality Protein Maize. In Proceedings of the 4th Annual Conference of the International Consortium on Agricultural Biotechnology, Ravello, Italy, 2000.
16. Kinoshita, T., Report of the committee on gene symbolization, nomenclature and linkage group. *Rice Genet. Newsl.*, 1995, **12**, 9–153.
17. Toenniessen, G. H., O'Toole, J. C. and DeVries, J., Advances in Plant Biotechnology and its adoptipon in developing countries. *Curr. Opin. Plant. Biol.*, 2003, **8**, 191–198.
18. Hittalmani, S., Parco, A., Mew, T. V., Ziegler, R. S. and Huang, N., Fine mapping and DNA marker assisted pyramiding of the three major genes for blast resistance in rice. *Theor. Appl. Genet.*, 2000, **100**, 1121–1128.
19. Beavis, N. D. and Keim, P., Identification of quantitative trait loci that are affected by environment. in Genoype-by-Environment Interaction. In *Genoype-by-Environment Interaction*, CRC Press, 1996, pp. 123–149.
20. Jiang, C. and Zeng, Z. B., Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics*, 1995, **140**, 1111–1127.
21. Tanksley, S. D. and Nelson, J. C., Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.*, 1996, **92**, 191–203.
22. Ribaut, J.-M. and Bertran, F. J., Single large scale marker assisted selection (SLS-MAS). *Mol. Breed.*, 1999, **5**, 531–541.
23. Sorrells, M. E. and Wilson, W. A., Direct classification and selection of superior alleles for crop improvement. *Crop Sci.*, 1997, **37**, 691–697.
24. Lynch, M. and Welsh, B., *Genetic Analysis of Quantitative Traits*, Sinauer Associates Inc., Sunderland, MA, USA, 1998.
25. Fridman, E., Pbleban, T. and Zamir, D., A recombination hotspot delimits a wild species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 4718–4723.
26. Frary, A. *et al.*, *fw2.2*: A quantitative trait locus key to the evolution of tomato fruit size. *Science*, 2000, **289**, 85–88.
27. Yano, M. *et al.*, *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice is closely related to the Arabidopsis flowering time gene CONSTANS. *Plant Cell*, 2000, **12**, 2473–2483.
28. Takahashi, Y., Shomura, A., Sasaki, T. and Yano, M., *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 7922–7927.
29. Stuber, C. W. *et al.*, Identification of genetic factors contributing to heterosis in a hybrid from elite maize inbred lines using molecular markers. *Genetics*, 1992, **132**, 823–829.
30. Stuber, C. W. *et al.*, Synergy of empirical breeding, marker assisted selection and genomics to increase crop yield potential. *Crop Sci.*, 1999, **39**, 1571–1583.
31. Tanksley, S. D. *et al.*, Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *Lycopersicon pimpinellifolium*. *Theor. Appl. Genet.*, 1996, **92**, 213–224.
32. Bernacchi, D. *et al.*, Advanced backcross QTL analysis of tomato: II. Evaluation of near isogenic lines carrying single donor introgression for desirable wild QTL alleles derived from *Lycopersicon hirsutum* and *Lycopersicon pimpinellifolium*. *Theor. Appl. Genet.*, 1998, **97**, 170–180.
33. Bernacchi, D. *et al.*, Advanced backcross QTL analysis of tomato: I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theor. Appl. Genet.*, 1998, **97**, 381–397.
34. Vincent, J. M. *et al.*, Marker assisted introgression of blackmold resistance QTL alleles from wild *Lycopersicon cheesmanii* to cultivated tomato (*Lycopersicon esculentum*) and evaluation of QTL phenotypic effects. *Mol. Breed.*, 2001, **8**, 217–233.
35. Ribaut, J.-M. *et al.*, QTL analyses, MAS results, and perspectives for drought tolerance improvement in tropical maize. In Workshop on Molecular Approaches for the Genetic Improvement of Cereals for Stable Production in Water-limited Environments, CIMMYT, Mexico, 2001.
36. Tanksley, S. D. *et al.*, RFLP mapping in plant breeding: new tools. *Bio/Technology*, 1989, **7**, 257–263.
37. Ribaut, J.-M. *et al.*, Simulation experiments on efficiencies of gene introgression by backcrossing. *Crop Sci.*, 2002, **42**, 557–565.
38. Gilles, G. *et al.*, Single nuclotide polymorphic discrimination by an electronic dot blot assay on semiconductor microchips. *Nature Biotechnol.*, 1999, **17**, 365–370.
39. Ribaut, J.-M. and Hoisington, D. A., Marker assisted selection: new tools and strategies. *Trends Plant Sci.*, 1998, **3**, 236–239.
40. Ribaut, J.-M. *et al.*, Genetic basis of physiological traits. In *Application of Physiology in Wheat Breeding*, CIMMYT, Mexico, 2001.
41. Ahn, S. N., Bollich, C. N. and Tanksley, S. D., RFLP analysis of genomic regions associated with cooked-kernel elongation in rice. *Theor. Appl. Genet.*, 1992, **84**, 825–828.
42. Peleman, J. D. and van der Voort, J. R., Breeding by design. *Trends Plant Sci.*, 2003, **8**, 330–334.
43. Zamir, D, Improving plant breeding with exotic genetic libraries. *Nature Rev. Genet.*, 2001, **2**, 983–988.
44. Stryer, L., Fluorescence energy transfer as a spectroscopic ruler. *Annu. Rev. Biochem.*, 1978, **47**, 819–846.
45. Livak, K. J., Flood, S. J. A., Marmaro, J., Giusti, W. and Deetz, K., Oligonucleotides with fluorescent dyes at opposite ends provide a

quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl.*, 1995, **4**, 357–362.
46. Salvi, S. R. *et al.*, Development of PCR based assays for allelic discrimination in maize by using 5′ nuclease procedure. *Mol. Breed.*, 2001, **8**, 169–176.
47. Meksem, K. *et al.*, High throughput genotyping for a polymorphism linked to soybean cyst nematode resistance gene *Rhg4* by using Taqman probes. *Mol. Breed.*, 2001, **7**, 63–71.
48. Landegren, U. R., Kaiser, R., Sanders, J. and Hood, L., Reading bits of genetic information: Methods for single nucleotide polymorphism analysis. *Science*, 1988, **241**, 1077–1080.
49. Tobe, V. O., Taylor, S. L. and Nickerson, D. A., Single well genotyping of diallelic sequence variations by a two-color ELISA-based ologonucleotide ligation assay. *Nucleic Acids Res.*, 1996, **24**, 3728–3732.
50. Chen, X. N., Livak, K. J. and Kwok, P. Y., A homogeneous, ligase mediated DNA diagnostic test for genome analysis. *Genome Res.*, 1998, **8**, 549–556.
51. Howell, W. M., Jobs, M., Gyllensten, U. and Brookes, A. J., Dynamic allele-specific hybridization. *Nature Biotechnol.*, 1999, **17**, 87–88.
52. Griffin, T. J., Hall, J. G., Prudent, J. R. and Smith, L. M., Direct genetic analysis by matrix assisted laser desorption/ionization mass spectrometry. *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 6301–6306.
53. Young, N. D., A cautiously optimistic vision for marker assisted breeding. *Mol. Breed.*, 1999, **5**, 505–510.
54. Heiter, P. and Boguski, M., Functional genomics: it's all how you read it. *Science*, 1997, **278**, 601–602.
55. Chu, S. *et al.*, The transcriptonal program of sporulation in budding yeast. *Science*, 1998, **282**, 699–705.
56. Wang, G. L. *et al.*, RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. *Genetics*, 1994, **136**, 1421–1434.
57. Liu, G., Lu, G., Zeng, L. and Wang, G. L., Two broad-spectrum blast resistance genes, *Pi9(t)* and *Pi2(t)* are physically linked on rice chromosome 6. *Mol. Genet. Genomics*, 2002, **267**, 472–480.
58. Chen, D. H. *et al.*, Molecular mapping of the blast resistance gene, *Pi44(t)* in a line derived from a durably resistant rice cultivar. *Theor. Appl. Genet.*, 1999, **98**, 1046–1053.
59. Yoshimura, S. *et al.*, RFLP analysis of introgressed chromosomal segments in three near isogenic lines of rice for bacterial blight resistance genes, xa-1, xa-3 and xa-8. *Jpn. J. Genet.*, 1992, **67**, 29–37.
60. Yoshimura, S. *et al.*, Tagging and combining bacterial blight resistance genes in rice using RAPD and RFLP markers. *Mol. Breed.*, 1995, **1**, 375–387.
61. Zhang, G., Angeles, E. R., Abenes, M. L. P., Khush, G. S. and Huang, N., RAPD and RFLP mapping of the bacterial blight resistance gene *xa*-13 in rice. *Theor. Appl. Genet.*, 1996, **93**, 65–70.
62. Ronald, P. C. *et al.*, Genetic and physical analysis of rice bacterial blight resistant locus, *xa-21. Mol. Gen. Genet.*, 1992, **236**, 113–120.
63. Lin, X. H., Zhang, D. P., Xie, Y. F., Gao, H. P. and Zhang, Q. F., Identifying and mapping a new gene for bacterial blight resistance in rice based on RFLP markers. *Phytopathology*, 1996, **86**, 1156–1159.
64. Sebastian, L. S., Ikeda, R., Huang, N., Imbe, T., Coffman, W. R. and McCouch, S. R., Genetic mapping of resistance to rice tungro spherical virus (RTSV) and green leaf hopper in ARC 11554. *Phytopathology*, 1996, **86**, 25–30.
65. Mohan, M. *et al.*, RFLP and RAPD mapping of the rice *Gm2* gene that confers resistance to biotype1 of gall midge. *Theor. Appl. Genet.*, 1994, **87**, 782–788.
66. Mohan, M. *et al.*, Molecular mapping of a resistance-specific PCR based marker linked to a gall midge resistance gene (Gm4t) in rice. *Theor. Appl. Genet.*, 1997, **95**, 777–782.
67. Sardesai, N. *et al.*, Identification and mapping of an AFLP marker linked to Gm 7, a gall midge resistance gene and its conversion to a SCAR marker for its utility in MAS in rice. *Theor. Appl. Genet.*, 2002, **105**, 691–698.
68. Hirabayashi, H. and Ogava, T., RFLP mapping of *Bph-1* in rice. *Breed. Sci.*, 1995, **45**, 369–371.
69. Mei, M., Zhuang, C., Wan, R., Wu, J., Hu, W. and Kochert, G., Genetic analysis and tagging of gene for brown plant hopper resistance in indica rice. In Proceedings of the III International Rice Genetics Symposium, Manila, Philippines, 1995, pp. 590–595.
70. Tamura, K. *et al.*, Mapping of the Grh1 locus for green rice leaf hopper resistance in rice using RFLP markers. *Breed. Sci.*, 1999, **49**, 11–14.
71. Xu, K. and Mackill, D. J., A major locus for submergence tolerance mapped on rice chromosome 9. *Mol. Breed.*, 1996, **2**, 219–224.
72. Zhang, J. S., Xie, C., Li, Z. Y. and Chen, S. Y., Expression of the plasma membrane H+ – ATPase gene in response to salt stress in a rice salt tolerant mutant and its original variety. *Theor. Appl. Genet.*, 1999, **99**, 1006–1011.
73. Yanagihara, S. *et al.*, Molecular analysis of the inheritance of the S-5 locus conferring wide compatibility in indica X japonica hybrids of rice. *Theor. Appl. Genet.*, 1995, **90**, 182–188.
74. Reddy, O. U. K. *et al.*, Genetic analysis of temperature sensitive male sterility in rice. *Theor. Appl. Genet.*, 2000, **100**, 794–801.
75. Wang, A. Y., Yu, W. P., Juang, R. H., Huang, J. W., Sung, H. Y. and Su, J. C., Presence of three rice sucrose synthase genes as revealed by cloning and sequencing of cDNA. *Plant Mol. Biol.*, 1992, **18**, 1191–1194.
76. Mackill, D. J., Salam, M. A., Wang, Z. Y. and Tanksley, S. D., A major photoperiod sensitive gene tagged with RFLP and isozyme marker in rice. *Theor. Appl. Genet.*, 1993, **85**, 536–540.
77. Liang, C. Z., Gu, M. H., Pan, X. B., Liang, G. H. and Zu, L. H., RFLP tagging of a new semidwarfing gene in rice. *Theor. Appl. Genet.*, 1994, **88**, 898–900.
78. Fukuta, Y. and Yagi, T., Mapping of a shattering resistance gene in a mutant line SR-5 induced from an indica rice variety, Nanjing 11. *Breed. Sci.*, 1998, **48**, 345–348.
79. Sobrizal, Ikeda, K., Sanchez, P. L. and Yoshimura, A., RFLP mapping of a seed shattering gene on chromosome 4 in rice. *Rice Genet. Newsl.*, 1999, **16**, 74–75.
80. Simcox, K. D. and Bennetzen, J. L., Clustering of disease resistance loci in the maize genome. *Phytopathology*, 1993, **83**, 1326–1330.
81. Liu, Z., Peter, S. O., Long, M., Weingartner, U., Stamp, P. and Kaeser, O., APCR assay for rapid discrimination of sterile cytoplasm types in maize. *Crop Sci.*, 2002, **42**, 566–569.
82. Maize GDB; http://www.maizegdb.org
83. Schachermayr, G. M., Messmer, M. M., Feuillet, C., Winzeler, H., Winzeler, M. and Keller, B., Identification of molecular marker linked to the Agropyron elongatum-derived leaf rust resistance gene *Lr 24* in wheat. *Theor. Appl. Genet.*, 1995, **90**, 982–990.
84. Gold, J., Harder, D., Townley-Smith, F. and Procunier, J., Development of a molecular marker for rust resistance genes *sr29* and *Lr35* in wheat breeding lines. *Electron. J. Biotechnol.*, 1999, **2**.
85. Schachermayr, G. M., Siedler, H., Gale, M. D., Winzeler, H., Winzeler, M. and Keller, B., Identification and localization of molecular markers linked to the *Lr 9* leaf rust resistance gene of wheat. *Theor. Appl. Genet.*, 1994, **88**, 110–115.
86. Helgurea, M., Khan, L. A. and Dubcovsky, J., Development of PCR markers for wheat leaf rust resistance gene *Lr 47. Theor. Appl. Genet.*, 2000, **101**, 625–631.
87. Hartl, L., Wiess, H., Stephan, U., Zeller, F. J. and Jahoor, A., Molecular identification of powdery mildew resistance genes in common wheat (*Triticum aestivum*). *Theor. Appl. Genet.*, 1995, **90**, 601–606.
88. Chantret *et al.*, Location and mapping of powdery mildew resistance gene MIRE and detection of a resistant QTL by BSA with microsatellites in wheat. *Theor. Appl. Genet.*, 2000, **100**, 1217–1224.

89. Dweikat, I., Zhang, W. and Ohm, H., Development of STS markers linked to hessian fly resistance gene H6 in wheat. *Theor. Appl. Genet.*, 2002, **105**, 766–770.
90. Williams, K. J. *et al.*, Mapping of root lesion nematode resistance gene *Ren1* in wheat. *Theor. Appl. Genet.*, 2002, **104**, 874–879.
91. Bullrich, L. *et al.*, Mapping of a thermo sensitive earliness *per se* gene *on Triticum monococcum* chromosome 1 A^m. *Theor. Appl. Genet.*, 2002, **105**, 585–593.
92. Procunier, J. D. *et al.*, DNA markers linked to a T10 loose smut resistance gene in wheat. *Genome*, 1996, **40**, 176–179.
93. Iwaki, K. *et al.*, Genetic analysis of *Vrn B-1* for vernalization requirement by using linked CAPS markers in bread wheat. *Theor. Appl. Genet.*, 2002, **104**, 571–576.
94. Khlestkina, E. K. *et al.*, Molecular mapping, phenotypic expression and geographical distribution of genes determining anthocyanin pigmentation of coleoptiles of wheat. *Theor. Appl. Genet.*, 2002, **104**, 632–637.
95. Parker, G. D. and Langridge, P., Development of a STS marker linked to a major locus controlling flour color in wheat. *Mol. Breed.*, 2000, **6**, 169–174.
96. Oh, B. J., Frederiksen, R. A. and Magill, C. W., Identification of molecular markers linked to head smut resistance gene (*shs*) in sorghum by RFLP and RAPD analyses. *Phytopathology*, 1994, **84**, 830–833.
97. Wen, L. *et al.*, Development and mapping of AFLP markers linked to the sorghum fertility restorer. *Theor. Appl. Genet.*, 2002, **104**, 577–585.
98. Yu, Y. G., Saghai Maroof, M. A., Buss, G. R., Maughan, P. J. and Tolin, S. A., RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. *Phytopathology*, 1994, **84**, 60–64.
99. Brummer, E. C., Nickell, A. D., Wilcox, J. R. and Shoemaker, R. C., Mapping the *Fan* locus controlling linolenic acid content in soybean oil. *J. Hered.*, 1995, **86**, 245–247.
100. Landau, E. D., Angermuller, S., Shoemaker, R. and Gresshoff, P. M., The genetic locus controlling supernodulation in soybean cosegregates tightly with a cloned molecular marker. *Mol. Gen. Genet.*, 1991, **228**, 221–226.
101. Rajesh, P. N. *et al.*, Identification of an STMS marker for the double podding gene in chickpea. *Theor. Appl. Genet.*, 2002, **104**, 604–607.
102. Schneider, A. *et al.*, Mapping of the nodulation loci *sym9* and *sym10* of pea. *Theor. Appl. Genet.*, 2002, **104**, 1312–1316.
103. Timmerman, G. M. *et al.*, Linkage mapping of *er*, a recessive *Pisum sativum* gene for resistance to powdery mildew fungus (*Erysiphe pisi*). *Theor. Appl. Genet.*, 1994, **88**, 1050–1055.
104. Tiwari, K. R., Penner, G. A. and Warkentin, T. D., Identification of coupling and repulsion phase RAPD markers for powdery mildew resistance gene *er*-1 in pea. *Genome*, 1998, **41**, 440–444.
105. Dirlewanger, E. P. *et al.*, Restriction Fragment Length Polymorphism analysis of loci associated with disease resistance genes and developmental traits in *Pisum sativum*. *Theor. Appl. Genet.*, 1994, **88**, 17–24.
106. Bouzidi, M. F. *et al.*, Molecular analysis of a major locus for resistance to downy mildew in sunflower with specific PCR-based markers. *Theor. Appl. Genet.*, 2002, **104**, 592–600.
107. Williamson, V. M., Ho, J. Y., Wu, F. F., Miller, N. and Kaloshian, I., A PCR based marker tightly linked to the nematode resistance gene, *Mi*, in tomato. *Theor. Appl. Genet.*, 1994, **87**, 757–763.
108. Yaghoobi, J., Kaloshian, I., Wen, Y. and Williamson, V. M., Mapping a new nematode resistance locus in *Lycopersicon peruvianum*. *Theor. Appl. Genet.*, 1995, **89**, 541–544.
109. Zamir, D. *et al.*, Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, *Ty-1*. *Theor. Appl. Genet.*, 1994, **88**, 141–146.
110. Robert, V. J. M. *et al.*, Marker assisted introgression of black mold resistance QTL alleles from wild *Lycopersicon cheesmanii* to cultivated tomato (*L. esculentum*) and evaluation of QTL phenotypic effects. *Mol. Breed.*, 2001, **8**, 217–233.
111. Gebhardt, C., Mugniery, D., Ritter, E., Salamini, F. and Bonnel, E., Identification of RFLP markers closely linked to the *H1* gene conferring resistance to *Globodera rostochiensis* in potato. *Theor. Appl. Genet.*, 1993, **85**, 541–544.
112. Marano, M. R., Malcuit, I., De Jong, W. and Baulcombe, D. C., High resolution genetic map of *Nb*, a gene that confers hypersensitive resistance to potato virus X in *Solnaum tuberosum*. *Theor. Appl. Genet.*, 2002, **105**, 192–200.
113. El-Kharbotly, A. *et al.*, Segregation analysis and RFLP mapping of the *R1* and *R3* alleles conferring race specific resistance to *Phytophthora infestans* in progenies of dihaploid potato parents. *Mol. Gen. Genet.*, 1994, **242**, 749–754.

Received 23 December 2003; revised accepted 2 June 2004