Genetic Markers, Map Construction, and Their Application in Plant Breeding

Jack E. Staub[1] and Felix C. Serquen[2]
Vegetable Crops Research, U. S. Department of Agriculture, Agricultural Research Service, Department of Horticulture, University of Wisconsin–Madison, WI 53706

Manju Gupta[3]
Mycogen Plant Sciences, Madison Laboratories, 5649 East Buckeye Road, Madison, WI 53716

The genetic improvement of a species through artificial selection depends on the ability to capitalize on genetic effects that can be distinguished from environmental effects. Phenotypic selection based on traits that are conditioned by additive allelic effects can produce dramatic, economically important changes in breeding populations. Genetic markers—heritable entities that are associated with economically important traits—can be used by plant breeders as selection tools (Beckman and Soller, 1983; Darvasi and Soller, 1994). Marker-assisted selection (MAS) provides a potential for increasing selection efficiency by allowing for earlier selection and reducing plant population size used during selection. Nevertheless, the phenotypic variation that marker loci define is often nonadditive, and is a function of genetic linkage, pleiotropy, and environment (Lark et al., 1995). Thus, the efficiency of application of marker loci as predictors of phenotypic variation depends on many factors, and predictions of response to selection (R) or genetic gain (Δ G) are often difficult.

The predictive value of genetic markers used in MAS depends on their inherent repeatability (Weeden et al., 1992), map position, and linkage with economically important traits (quantitative or qualitative). The presence of a tight linkage (<10 cM [centimorgan]) between qualitative trait(s) and a genetic marker(s) may be useful in MAS to increase gain from selection (Kennard et al., 1994; Paran et al., 1991; Timmerman et al., 1994). Likewise, selection for multiple loci or quantitative trait loci (QTL) using genetic markers can be effective if a significant association is found between a quantitative trait and markers (Edwards and Page, 1994; Edwards et al., 1987; Lande and Thompson, 1990).

Often the biotechnological information presented in research reports is not tied directly to classical genetic methodologies and the sophisticated technology presented results in a bewildering array of new terms. For scientists who have a peripheral interest in genome mapping, but would like to understand the potential role of MAS in plant improvement, the wealth of information currently being produced in this area can lead to considerable confusion. The purpose of this paper is to describe available marker types and examine factors critical for their use in map construction and MAS. This review clarifies how genetic markers are used in map construction and defines the potential use of genetic maps for MAS.

Received for publication 14 July 1995. Accepted for publication 24 Nov. 1995. Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked *advertisement* solely to indicate this fact.

[1]Research Horticulturist and Professor.
[2]Research Associate.
[3]Research Scientist, Applied Genetics.

MARKER TYPES

Morphological. Morphological traits controlled by a single locus can be used as genetic markers if their expression is reproducible over a range of environments. Although codominant morphological markers have been useful as predictors of genetic response to selection, they can be influenced by environmental and genetic factors (e.g., epistasis). For instance, the expression of the determinate (*de*) character in cucumber (*Cucumis sativus* L.) may vary, depending on growing environment and modifying genes (Staub and Crubaugh, 1995). Thus, a description of such a trait has significance only when accompanied by properly documented pedigree information and environmental conditions. The fact that such factors may modify a gene's expression of phenotype may limit its usefulness as a genetic marker. A further drawback of morphological markers is that they may present an altered phenotype that interferes with grower needs.

Isozymes. Isozymes are differently charged protein molecules that can be separated using electrophoretic procedures (usually starch gel) (Markert and Moller, 1959). Since enzymes catalyze specific biochemical reactions, it is possible to visualize the location of a particular enzyme on a gel by supplying the appropriate substrate and cofactors, and involving the product of the enzymatic reaction in a color-producing reaction. The colored product becomes deposited on the gel, forming a visible band where a particular enzyme has been electrophoretically localized. Bands visualized from specific enzymes represent protein products, have a genetic basis, and can provide genetic information as codominant markers. However, the paucity of isozyme loci and the fact that they are subject to post-translational modifications often restricts their utility (Staub et al., 1982).

RFLPs. Restriction fragment length polymorphisms (RFLPs) are detected by the use of restriction enzymes that cut genomic DNA molecules at specific nucleotide sequences (restriction sites), thereby yielding variable-size DNA fragments (Fig. 1). Identification of genomic DNA fragments is made by Southern blotting, a procedure whereby DNA fragments, separated by electrophoresis, are transferred to nitrocellulose or nylon filter (Southern, 1975). Filter-immobilized DNA is allowed to hybridize to radioactively labeled probe DNA. Probes are usually small [500 to 3000 base pairs (bp)], cloned DNA segments (e.g., genomic or cDNA). The filter is placed against photographic film, where radioactive disintegrations from the probe result in visible bands. Such bands are visualizations of RFLPs, which are codominant markers.

The polymerase chain reaction (PCR) has been used to develop several DNA marker systems (Fig. 1). Three strategies primarily have been employed in the development of PCR-based marker systems. These include: 1) markers that are amplified using single primers in PCR, where marker system diversity results from variation in the length and/or sequence of primers, and where anchor nucleotides are present at 5´ or 3´ termini of primers (e.g., RAPDs, SPARs, DAFs, AP-PCR, SSR-anchored PCR; see below); 2) markers that are selectively amplified with two primers in PCR such that their selectivity comes from the presence of two to four random bases at the 3´ ends of primers that anneal to the target DNA during the PCR (e.g., AFLP, below); and 3) markers amplified using two primers in PCR, which commonly requires cloning and/or sequencing for the construction of specific primers. In this case, variations in marker technology result from differences in the target DNA sequence present between two primers (e.g., AMP-FLPs, STRs, and SSRs).

RAPDs. Of three similar, single-primer, PCR-based technologies, random amplified polymorphic DNA (RAPD; Williams et al., 1990), DNA amplification fingerprinting (DAF; Caetano-Anolles et al., 1991), and arbitrary-primed PCR, (AP-PCR; Owen and Uyeda, 1991; Welsh and McClelland, 1990), RAPDs have been used most widely for map construction and linkage analysis (Reiter et al., 1992) (Fig. 1). RAPD markers are generated by PCR amplification of random genomic DNA segments with single primers [usually 10 nucleotides (nt) long] of arbitrary sequence (Williams et al., 1990). The primer/target complexes are used as substrates for

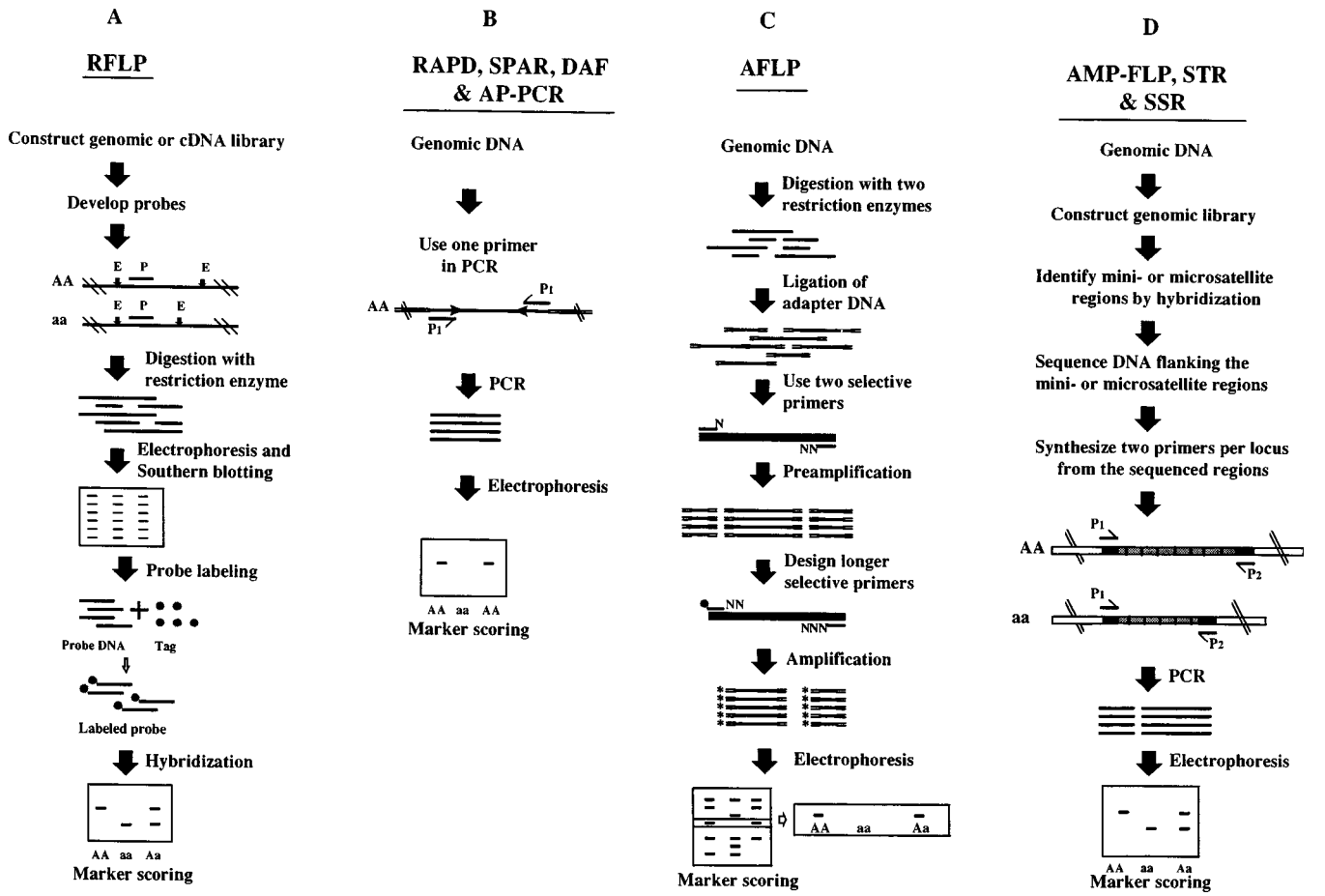

Fig. 1. Comparisons of various molecular marker technologies. In contrast to technologies presented in panels **B** and **C**, those in panels **A** and **D** require an expensive and time-consuming set-up phase for probe and primer development. Based on their availability, probes and primers can also be obtained from institutions or universities. DNA markers in panels **B** and **C** can be generated within a short time and require minor development. (**A**) E = restriction enzyme, P = probe, tag = either radioactive or nonradioactive for detection of hybridized probes. (**B**) P_1 = single primer of polymerase chain reaction (PCR). Although most PCR-based technologies amplify few to multiple bands, examples of only one band are given to illustrate scoring. (**C**) Adapter DNA = short double-stranded DNA molecules, 18 to 20 base pairs in length, representing a mixture of two types of molecules. Each type is compatible with one E generated DNA ends. Selective primers = primers for preamplification designed to contain an adapter DNA sequence plus one to two random bases at the 3′ end for reading into the genomic fragments. Primers for amplification have the preamplification primer sequence plus one to two additional bases at the 3′ end. A tag is attached at the 5′ end of one of the amplification primers for detecting amplified molecules (multiple bands are also possible). Differences in scoring exist between bands in this panel and those of panel **B** (also see Fig. 2). (**D**) P1 and P2 = two PCR primers.

DNA polymerase to copy the genomic sequences 3′ to the primers. Iteration of this process yields a discrete set of amplified DNA products that represent target sequences flanked by opposite-oriented primer annealing sites. Amplification products can be separated by electrophoresis on agarose or polyacrylamide gels and visualized by staining with ethidium bromide or silver. RAPDs are usually dominant markers with polymorphisms between individuals defined as the presence or absence of a particular RAPD band (Fig. 2). The further development of RAPD methodology has produced other PRC-based markers (e.g., SCAR and ASAP markers).

SCARs. Utility of a desired RAPD marker can be increased by sequencing its termini and designing longer primers (e.g., 24 nt) for specific amplification of markers (Paran and Michelmore, 1993). Such sequenced characterized amplified regions (SCARs) are similar to sequence-tagged sites (STS) (Olson et al., 1989) in construction and application. DNA sequence differences are manifest by the presence or absence of a single unique band. SCARs are more reproducible than RAPDs and can be developed into plus/minus arrays where electrophoresis is not needed. Although SCARs are usually dominant markers, some SCARs can theoretically be converted to codominant markers by digestion with 4-bp restriction endonucleases, and identification of polymorphisms by either denaturing gradient gel electrophoresis (DGGE) or single-strand conformational polymorphism (SSCP) techniques (Rafalski and Tingey, 1993).

ASAPs. A recent modification of PCR technology involves the alkaline extraction of DNA with subsequent amplification of the DNA template in microtiter plates using allele-specific associated primers (ASAPs) that generate only a single DNA fragment at stringent annealing temperatures (similar to SCARs) (Gu et al., 1995). The DNA fragment is present in only those individuals possessing the appropriate allele and thus eliminates the need to separate amplified DNA fragments by electrophoresis (i.e., presence/absence polymorphism). This method involves ethidium bromide binding to the DNA double helix, which dramatically enhances its fluorescence but does not bind to free nucleotides in the PCR mixture. This approach was developed to decrease time for DNA extraction and increase the reliability of the PCR reaction for large-scale screening.

SPARs. The single primer amplification reaction (SPAR) is a DNA marker system that can produce multiple markers per assay (Fig. 1). The system uses primers based on microsatellites or simple sequence repeats (SSRs) and amplifies inter-SSR DNA sequences (Gupta et al., 1994). Of the di-, tri-, tetra-, and pentanucleotide SSRs, the tetranucleotide repeats are most effective in producing polymorphic multiband patterns. The level of polymorphism is related to the genomic diversity within a given species. Most DNA markers map to scattered genomic locations. Although most SSR-SPARs are dominant markers, codominant markers can also be detected (Fig. 2). Given that an unlimited number of primers can be synthesized from the tetranucleotide repeats [$(4)^4 = 256$], and from the combination of di-, tri-, and

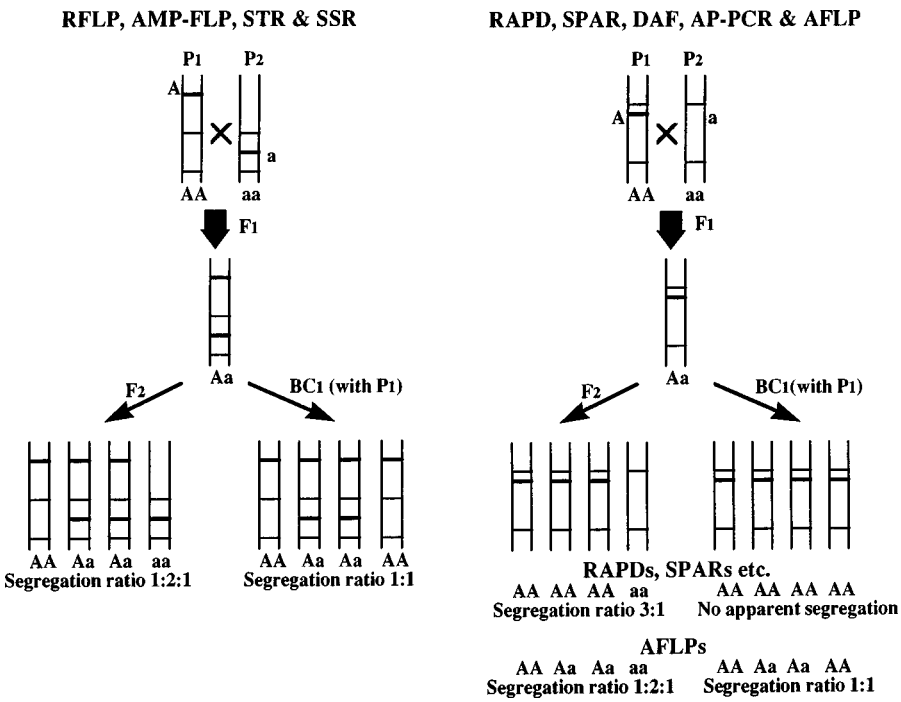

Fig. 2. Schematic of inheritance patterns of DNA markers in F_2 and BC_1 progenies. Common bands shown in the patterns of the parents and progeny may not be seen with other probes and primers. RFLPs, AMP-FLPs, STRs, and SSRs are codominant markers and thus heterozygousity in F_2 and BC_1 is easily detected. In contrast, RAPDs, SPARs, DAPs, AP-PCR, and AFLPs are dominant markers and detection of heterozygosity is, with rare exception, not possible. Zygosity determination is also possible through the quantification of DNA bands (e.g., AFLPs).

tetranucleotide SSRs, or compound SSRs, the SSR-SPAR marker system may have broad application across a range of plant species.

SSR-anchored PCR. This system employs single primers of dinucleotide simple sequence repeats (SSRs; see below); especially (CA)n repeats for amplification of markers. The primer is either anchored at 3′ or 5′ termini with two to four nucleotides (Zietkiewicz et al., 1994). Multiple bands containing inter-SSR regions are amplified and then are fractionated on polyacrylamide gels for pattern visualization. These amplified bands are mostly dominant markers and can be used in a wide range of plant species.

AFLPs. Production of amplified fragment length polymorphisms (AFLPs) is based on selective amplification of restriction enzyme-digested DNA fragments (Zabeau and Vos, 1993) (Fig. 1). Multiple bands are generated in each amplification reaction that contains DNA markers of random origin. Analysis of DNA on denaturing polyacrylamide gels typically results in the production of 50 to 100 bands per individual sample. AFLPs are quantitative in that heterozygous and homozygous genotypes can be differentiated by the intensity of the amplified bands. The ability of this technology to generate many markers with minimum primer testing, and the system's high resolution (i.e., band clarity and relatively low lane background) are features that make AFLPs attractive as genetic markers (primarily dominant; Fig. 2). Because of its expense, automation may be required to realize this technology's full potential during MAS.

AMP-FLPs, STRs, and SSRs. Mini- and microsatellite DNA sequences are an excellent source of polymorphisms in eukaryotic genomes, and are well suited for genotyping and map construction. Marker systems based on such sequences include amplified fragment length polymorphisms (AMP-FLPs; minisatillites in vertebrates; Fregeau and Fourney, 1993), short tandem repeats (STRs; microsatellites in verebrates; Fregeau and Fourney, 1993), and simple sequence repeats (SSRs; microsatellites in plants; Rafalski and Tingey, 1993) (Fig. 1). Mini- and microsatellites are comprised of tandem arrays of 15- to 70-bp and 2- to 5-bp monomeric repeat units, respectively. Polymorphisms appear because of variation in the number of tandem repeats in a given repeat motif. Most STRs and SSRs are dinucleotide repeat-based [(AC)n, (AG)n, and (AT)n] microsatellite markers (Rafalski and Tingey, 1993). Such polymorphisms are amplified by designing primers from the sequenced regions flanking the repeat motifs (Fig. 1). Similar to (CA)n repeats in humans, (AT)n dinucleotide microsatellite repeats are relatively abundant and highly polymorphic in plants (Akkaya et al., 1992; Morgante et al., 1994). This group of markers is codominant in its expression (Fig. 2).

CAPs. Cleaved amplified polymorphic sequences (CAPs) are a form of genetic variation in the length of DNA fragments generated by the restriction digestion of PCR products (Koniecyzn and Ausubel, 1993; Jarvis et al., 1994). The source of the sequence information for the primers can come from a genebank, genomic or cDNA clones, or cloned RAPD bands. This marker class is codominant in its behavior.

MARKER SYSTEM SELECTION

Selection of a DNA marker system for plant breeding depends on project objectives, population structure, the genomic diversity of the species under investigation, marker system availability, time required for analysis, and the cost per unit information (Table 1). Clearly, each marker system has advantages and disadvantages, and therefore it is critical to evaluate each marker system for its potential utility before use. For example, intraspecific maps can be constructed with a common set of RFLP probes; however, each species initially requires the construction of a map. The range of polymorphism in species also plays a role in marker selection (e.g., in self-pollinated species RAPDs are more useful for detecting polymorphisms within a gene pool than RFLPs). Moreover, the use of a marker system in one species does not necessarily indicate its efficacy in another species.

Marker systems also differ in their utility across populations, species, and genera, and their efficacy in the detection of polymorphism. For instance, RFLPs mapped in one population can be used as probes for characterizing other populations within the same species. In contrast, SSRs can be as informative as RFLPs, but polymorphic primers identified in one species are generally not useful in another species. Likewise, maps using RAPDs, SPARs, and AFLPs can be constructed in a relatively short period; however, such markers are not useable across populations, because each marker is primarily defined by its length (i.e., sequence information may be limited). Moreover, the same size band amplified across populations/species does not necessarily mean that bands possess the same sequence, unless proven by hybridization studies (Thormann et al., 1994). In contrast to RFLPs, these marker systems possess all the advantages of PCR-based systems (i.e., small sample requirement, high throughput, and early selection) (Table 1). These advantages can be nullified if polymorphism within a species is low. Where the level of polymorphism is low, STRs and SSRs are currently the marker systems of choice. However, the cost and time required to develop these marker systems can be considerable (Table 1).

Costs per unit information (data point) depend on the time required for DNA extraction (sampling), the amount of DNA needed for analysis, whether cloning and sequencing is necessary, the amount of potentially useful genetic information acquired, the type of genetic information needed, whether the allelic variation can be ascribed to banding patterns (dominant vs. codominant), whether the electrophoretic system can be automated, the potential utility of genetic maps, and the proprietary status of the technique (Table 1). Codominant markers, such as RFLPs, are useful for MAS and evolutionary studies, but their use can be time consuming, relatively expensive, and may require considerable technical exper-

Table 1. Comparisons among several molecular marker systems for various technical attributes and proprietary rights status.

	Molecular marker systems[z]					
Critical variables	RFLP	AFLP	SPAR	RAPD	DAF	SSR/STR/AMP-FLP
Tissue sampling (weeks)[y]	4	2	2	2	2	2
DNA needed/100 markers (mg)	100	0.5	2.5	2.5	2.5	10
Cloning and sequencing	Yes	No	No	No	No	Yes
Information content per run[x]	0–3	0–30	0–20	0–4	0–10	0–2
Marker type[w]	C	D	D	D	D	C
Zygosity detection[v]	Yes	Yes	No	No	No	Yes
Automation[u]	+	+++	++	++	++	+++
Utility of genetic maps[t]	SS	CS	CS	CS	CS	SS
Proprietary rights status[s]	NA	LC	NA	LC	LC	NA

[z]RFLP = restriction fragment length polymorphism, AFLP = amplified fragment length polymorphism, SPAR = single primer amplification reaction, RAPD = random amplied polymorphic DNA, DAF = DNA amplification fingerprinting, and SSR/AMP-FLP = simple sequence repeats/amplified fragment length polymorphism.
[y]The sampling time after sowing is shown for corn (*Zea mays*) (relative time applies to all crop species).
[x]Markers obtained per hybridization or PCR reaction. For SSR/AMP-FLP, number of markers per run does not reflect multiplexing.
[w]D and C equal dominant and codominant markers, respectively.
[v]Heterozygous alleles can be distinguished from the homozygous alleles.
[u]On the scale of 1 to 3, + = the least and +++ shows the most potential for automation. Automation refers to mechanizing steps involving processing of DNA and detection, identification and scoring of markers.
[t]Refers to the relative utility of maps constructed with a given marker system either within that species (SS = species specific) or to a specific cross or population (CS = cross specific).
[s]NA = not applicable, LC = license required to practice the technology.

tise. Often the high cost of developing an RFLP marker system for a new species or inefficiency of MAS, due to the large quantity of DNA required and slow screening process, results in a decision to use a PCR-based system. Nevertheless, RFLPs command an advantage in systematic and evolutionary studies because, in contrast to most PCR-based technologies, which detect variation in 20–40 bases (combined length of primers), data aquisition is based on the homology among large fragments of DNA (length of probes). The use of RFLPs for such studies is enhanced if polymorphisms are abundant.

MAP CONSTRUCTION

The development of molecular marker technology and consequent identification of many marker loci has caused renewed interest in genetic mapping. Genetic map construction requires that the researcher: 1) select the most appropriate mapping population(s); 2) calculate pairwise recombination frequencies using these population(s); 3) establish linkage groups and estimate map distances; and 4) determine map order. Since large mapping populations are often characterized by different marker systems, map construction has become computerized. Computer packages such as Linkage 1 (Suiter et al., 1983), GMendel (Echt et al., 1992), Mapmaker (Lander and Botstein, 1986; Lander et al., 1987), MapManager (Manly and Elliot, 1991), and JoinMap (Stam, 1993) have been developed to aid in the analysis of genetic data for map construction. These programs use data obtained from segregating populations to estimate recombination frequencies that are then used to determine the linear arrangement of genetic markers by minimizing recombination events.

Mapping populations

Selection of mapping populations is critical to successful map construction. Since a map's economic significance will depend upon marker-trait associations, as many qualitatively inherited morphological traits as possible should be included in genetic stocks chosen as parents for map construction (Table 2). Consideration must be given to the source of parents (adapted vs. exotic) used in the mapping population. Chromosome pairing and recombination rates can be severely disturbed (suppressed) in wide crosses (adapted x exotic) and generally yield greatly reduced linkage distances (Albini and Jones, 1987; Zamir and Tadmor, 1986). Wide crosses will usually provide segregating populations with a relatively large array of polymorphisms when compared to progeny segregating in a narrow cross (adapted x adapted). To have significant value in plant improvement programs, a map made from a wide cross must be colinear (i.e., order of loci similar) with maps constructed using adapted parents.

The choice of an appropriate mapping population depends on the type of marker systems employed (Tanksley et al., 1988). Maximum genetic information is obtained from a completely classified F_2 population using a codominant marker system (Mather, 1938). Information from a dominant marker system can be equivalent to a completely classified F_2 population if progeny tests (i.e., F_3 or F_2BC) are used to identify heterozygous F_2 individuals. This procedure is often prohibitive because of the cost and time involved in progeny testing.

Dominant markers supply as much information as codominant markers in recombinant inbred lines (RI) (i.e., an array of genetically related lines; usually $>F_8$), doubled haploids, or backcross populations in coupling phase (Burr et al., 1988). Information obtained from dominant markers can be maximized by using RI or doubled haploids because all loci are homozygous, or nearly so. Under conditions of tight linkage (i.e., about <10% recombination), dominant and codominant markers evaluated in RI populations provide more information per individual than either marker type in backcross populations (Reiter et al., 1992). However, as the distance between markers becomes larger (i.e., loci become more independent), the information obtained per unit individual in RI populations decreases dramatically when compared to codominant markers.

Backcross populations can be useful for mapping dominant markers if all loci in the recurrent parent are homozygous and the donor and recurrent parent have contrasting polymorphic marker alleles (Reiter et al., 1992). Information obtained from backcross populations using either codominant or dominant markers is less than that obtained from F_2 populations because one, rather than two, recombinant gametes are sampled per plant. Backcross populations, however, are more informative (at low marker saturation) when compared to RIs as the distance between linked loci increases in RI populations (i.e., about >15% recombination). Increased recombination can be beneficial for resolution of tight linkages, but may be undesirable in the construction of maps with low marker saturation.

Progeny testing of F_2 individuals is often used in map construction where phenotypes do not consistently reflect genotype (e.g., disease reaction and many useful traits) or where trait expression is controlled by QTLs. Segregation data from progeny test populations (e.g., F_3 or F_2BC) can be used in map construction. MAS can then be applied to cross progeny based on marker-trait map associations, especially in early generations (F_2, F_3), where linkage groups have not been completely disassociated by recombination events (i.e., maximum disequilibrium).

Recently, a method has been developed for the rapid identification of linkage using bulked segregant analysis (BSA; Michelmore et al., 1991). In BSA, two bulked DNA samples are drawn from a segregating population originating from a single cross. These bulks contain individuals that are identical for a particular trait (e.g., resistant or susceptible to a particu-

lar disease) or genomic region but arbitrary at all unlinked regions (i.e., heterozygous). The bulks are screened for DNA polymorphisms and these differences compared against a randomized genetic background of unlinked loci. Thus, differences between the two bulks indicate markers (e.g., bands) that are linked to a particular trait. BSA overcomes several problems that are associated with the use of nearly isogenic lines (NILs), which require many backcrosses to develop. Where only a portion of the polymorphic loci are expected to map to a selected region using NILs (e.g., BC_5 only 50%), regions unlinked to the target region will not differ between the bulked samples of many individuals in BSA. Moreover, all loci detected during BSA will segregate and can be mapped, thus eliminating the linkage drag problems (i.e., genes incorporated into lines by backcrossing that are flanked by DNA segments introduced from the donor parent) associated with NILs (Young and Tanksley, 1989).

Calculation of recombination fraction

Crossover events can be described as the percentage of recombination in offspring. Only half the meiotic products will be crossover types (recombinants) when one chiasma forms between two loci. Multiple crossovers can also be detected through observation of progeny phenotypes. Single crossover events are not independent and the number of double crossover events is usually smaller than predicted. This positive "interference" varies, depending on organism, crossover location, environmental factors, and numerous other factors. Therefore, accurate estimates of double crossing over can only be obtained when interference is considered. Interference is measured as a coefficient of coincidence (CC), which is an expression of the ratio of observed double crossovers to those predicted (expected) by a map. The expected double crossover frequency is predictable if two crossovers are independent events or if interference can be measured.

The proportion of mean number of recombination events defines the map distance between two loci. The relationship between map distance and recombination value is characterized by a genetic mapping function (*mf*). An *mf* is a formula expressing quantitative relationships between distances in a linkage map using crossover frequency. There are several types of mapping functions that can be applied, depending on the assumed degree of crossover interference that best represents the mapping population. The most common mapping functions were developed by Haldane (1919) and Kosambi (1944). While Haldane's *mf* assumes absence of interference, Kosambi's assumes positive interference (i.e., fewer double recombinants when compared to no interference).

The frequency of recombinant gametes produced can be used as an index of the distance between two loci on a chromosome [1 map unit = about 1 cM]. Map distance, however, is not completely additive. Additivity is based on the assumption that the average number of crossovers per chromatid occurring between two loci is directly proportional to the distance between the two loci. The frequency of recombination (percentage) and map distance is, however, not directly proportional. Estimates of the frequency of crossing over will be most reliable when genes are relatively closely linked (1 to 10 map units). Recombination percentage is only equivalent to map distance within the range of the minimum distance for crossing over (lack of additivity) because double crossovers occur at significant frequency. A nonlinear relationship occurs

Table 2. Total map length and mean distance between genetic markers in various plant species.

Common name	Botanical name	Marker Pop.[z]	Marker Type[y]	Marker No.	Map Length (cM)	Map Mean distance (cM)	Reference
Arabidopsis	*A. thaliana* L.	RI	RFLP	320	630	2.0	Reiter et al., 1992
Banana	*Musa acuminata* Colla	F_2	RFLP, isozyme, RAPD	90	606	10.0	Faure et al., 1993
Barley	*Hordeum vulgare* L.	DH	RFLP, isozyme, RAPD, morphological, SAP, disease	295	1250	4.2	Kleinhofs et al., 1993
Bean	*Phaseolus vulgaris* L.	BC	RFLP, isozyme, seed protein, flower color	244	1200	≈5.0	Vallejos et al., 1992
		F_2	RFLP, RAPD, isozyme	152	827	6.5	Nodari et al., 1993a
Brassica	*B. napus* L.	F_2	RFLP, isozyme, RAPD, morphology, disease	120	1413	14.0	Landry et al., 1991
	B. napus L.	F_1-DH	RFLP	132	1016	7.7	Ferreira et al., 1994
	B. rapa L. (syn. *campestris*)	F_2	RFLP	280	1850	6.9	Song et al., 1991
	B. rapa	F_3	RFLP, seed color, seed erucic acid, pubescence	139	1785	13.5	Teutonico & Osborn, 1994
Citrus	*C. grandis* L. x *C. paradisi* Maef.	F_1	RFLP	46	1700	20.0	Jarrell et al., 1992
Cotton	*Gossypium hirsutum* L. x *G. barbadense* L.	F_2	RFLP	705	4675	7.1	Reinisch et al., 1994
Cucumber	*Cucumis sativus* L.	F_2	RFLP, RAPD, isozyme, morphology, disease	58	766	8.0	Kennard et al., 1994
	C. sativus x *C. hardwickii* (R.) Alef.	F_2	RFLP, morphology	70	480	8.1	
Cuphea	*Cuphea lanceolata* Ait.	F_2	RFLP	37	288	7.8	Webb et al., 1992
Maize	*Zea mays* L.	RI	RFLP	334	1460	≈5.0	Burr et al., 1988
Potato	*Solanum tuberosum* L. x *S. berthaultii* Corr.	BC	RFLP, isozyme	977	684	0.7	Tanksley et al., 1992
Rice	*Oryza sativa* L. x *O. longistaminata* A. Chev. & Roehr	BC	RFLP	726	1491	2.0	Causse et al., 1994
Rye	*Secale cereale* L.	IBL	RFLP, isozyme, RAPD, morphology, physiology	60	350	6.0	Philipp et al., 1994
Sorghum	*S. bicolor* L.	F_2	RFLP	98	949	10.0	Whitkus et al., 1992
		F_2	RFLP	190	1789	9.4	Xu et al., 1994
Soybean	*Gylcine max* L.	F_2	RFLP	252	2147	8.5	Diers et al., 1992
Sugar beet	*Beta vulgaris* L.	F_2	RFLP	115	789	6.9	Pillen et al., 1992
Tomato	*Lycopersicon esculentum* Miller x *L. pennelli*	F_2	RFLP, isozyme	1030	1276	1.2	Tanksley et al., 1992

[z]RI = recombinant inbred, DH = doubled haploid, and IBL = inbred line.
[y]RFLP = restriction fragment length polymorphism, RAPD = random amplified polymorphic DNA, and SAP = specific amplicon polymorphism.

when estimates are made outside of this distance. Therefore, the actual map distance between two genes will tend to be underestimated by the recombination fraction (r) (e.g., $r = 0.10 = 10$ cMs, $r = 0.30 = 45.8$ cMs, $r = 0.35 = 60.2$ cMs), such that at large distances (~ 40–50 map units) the two genes will be strictly independent of each other (Kosambi, 1944).

Linkage phase

Genes are linked when they are on the same chromosome. There are two possible arrangements of two genes on a pair of chromosomes—coupling and repulsion. Coupling signifies that the two recessive alleles are carried in one chromosome and the two dominant alleles in the other (i.e., *AB/ab*); repulsion describes the alternate arrangement (i.e., *Ab/aB*). This relationship is particularly important when dealing with dominant markers during map construction. Two linked markers scorable as dominant alleles [e.g., *AA* or band presence (+)] can only be recognized in coupling phase linkage. This is because the heterozygote class cannot be distinguished from the homozygote dominant class (i.e., band presence = *AA* or *Aa*, band absence = *aa*). In contrast, codominant markers allow for the expression of both pairs of alleles (i.e., phenotypically as *AA, Aa, aa*). Linkage phase has proven important in selection when dominant markers are used (Haley et al., 1994). A greater proportion of bean (*Phaseolus vulgaris* L.) genotypes homozygous resistant to bean common mosaic virus (BCMV), and a lower proportion of segregating and homozygous susceptible genotypes were recovered when selection was imposed against a repulsion-phase RAPD marker than when selection was made for the coupling-phase RAPD marker. This observation is of practical significance where resistance is conditioned by recessive genes since it requires breeders to select against the heterozygous susceptible individuals (Kelly, 1995). In the case of selection for the recessive BCMV resistance gene (*bc-3*) in bean, Kelly stated: "selection of individuals based on the phenotype of combined coupling and repulsion-phase RAPD markers was equivalent to selection based on a codominant marker (RFLP) and was identical to selection based on the repulsion-phase marker alone."

Establishment of linkage groups

If linkage is indicated by Chi-square analysis of progeny segregation, then the potential for linkage between loci can be mathematically tested. There are several mathematical methods available for investigating potential linkage relationships (Crow, 1990). Among these are the maximum likelihood and least squares/regression methods. Currently, these are the methods of choice for linkage estimation because they result in estimates that have the smallest standard error (Mather, 1938; Nordheim et al., 1984). They are especially useful where multiple loci (QTL) are involved (Shute, 1988). While least squares estimation attempts to minimize deviations from a mathematical model (regression), maximum likelihood involves comparisons among two or more plausible hypotheses (e.g., linkage vs. no linkage). The maximum likelihood method is particularly useful in evaluating genetic phenomena and will be used in this discussion as an illustration of linkage analysis (Chakravarti et al., 1991).

Maximum likelihood and likelihood odds ratio (LOD) value

Mather (1938) developed the maximum likelihood approach for linkage analysis. It is used by various computer-based linkage programs (e.g., Mapmaker; Lander et al., 1987) to determine the probability of linkage between a given marker and a known marker. Maximum likelihood is a statistical procedure designed to choose values for variables that maximize a defined function, which is done by integrating the function and solving for 0 (minimizing the integrated function) or by iteration.

Linkage estimation using the method of maximum likelihood is based on the binomial expansion, which is a special case of the polynomial $(m_1 + m_2 + m_3 + \ldots)^n$. Maximum likelihood, as applied to linkage estimation, attempts to select a linkage estimator (r value) that minimizes an expectation function in a binomial expression. The benefit of maximum likelihood in the calculation and estimation of r is that functions can be designed that include ambiguous classes. An example of an ambiguous class is the double heterozygote (*AaBb*) of an F_2 family that contains recombinant and non-recombinant types.

Recombination value is used in the maximization expression to determine the likelihood (L) of association of linkage between a set of variables (i.e., genetic loci). The value of r that maximizes the likelihood of the observed outcome is determined. Solutions are limited to the range of 0 to 0.5.

After maximization, the question is raised as to whether the value of r, say x, is significant, given the upper limit of no linkage (i.e., 0.5)—that is, whether the probability that two loci are linked with a given r value over the probability that the two loci are not linked. An understanding of the precision of r is necessary to assess the utility of the value obtained by likelihood maximization. Historically, this has been done in two ways. Allard (1956) constructed a series of tables and formulae to calculate recombination values and associated standard errors using a maximum likelihood approach. This approach had already been widely used by plant geneticists (Fisher, 1946; Kramer and Burnham, 1947). Researchers in human genetics gravitated toward the use of the LOD defined by Haldane and Smith (1947). This approach has been used by some plant researchers because the LOD calculations needed for analyzing large populations and using many markers has been simplified by the use of some computer-based linkage estimation programs (e.g., Mapmaker; Lander and Botstein, 1986, Lander et al., 1987).

The odds ratio of a maximization event is given as: L (x)/L (0.5). This form, however, is inconvenient in most instances and the log of the odds ratio {i.e., LOD = log [L (x)/L (0.5)]} is used (Risch, 1992). In many analyses, a significance level of LOD > 3.0 is appropriate as an acceptance level of linkage between two loci. This value is equivalent to saying that the alternative hypothesis (linkage) has to be greater than 1000 times more probable than the null (no linkage) hypothesis. If this analysis is repeated over 100 marker loci, a significant level of LOD > 3 for each locus is comparable to an experiment-wise (genome-wise) type I error rate of alpha (α) = 0.01. LOD decreases with increasing r values and increases with increasing sample size.

Tests of linkage for qualitatively inherited traits vary in scope and operation. The researcher must determine a threshold LOD value below which linkage is not considered significant (Churchill and Doerge, 1994). As the LOD threshold is raised, fewer markers are assigned to linkage groups (i.e., independent loci), and more and smaller linkage groups are identified. Comparison of maps created from an array of LOD values often allows the researcher to determine the stability of putative linkage groupings. It is clear that any map only approximates reality and that map distances between markers will change as new information (i.e., more markers) becomes available.

Gene order determination and map merging

Because additivity of map distance is accepted (assuming no double crossovers occur) for narrow intervals (1–10 map units), tightly linked genes can be placed in relative order. Genes that are loosely linked (>20 map units) can be placed on a map but their location is much more tentative. The map distances calculated based on crossover percentages (i.e., genetic map) often bear no direct relationship with the actual physical distances between linked genes (i.e., physical map) (Stansfield, 1969; Swanson et al., 1990). The linear order in the physical and genetic maps, however, should theoretically be identical.

Three linked genes may be in any one of three orders, depending on which gene is in the middle of the linkage group. Traditionally, gene orders have been determined from either two- or three-point testcross data. When multiple crossovers occur with much greater than random frequency (i.e., localized negative interference), gene order of closely linked sites can be ascertained using three-factor reciprocal crosses.

Genetic maps in several crop species have been constructed using various marker systems, types of populations, and, often, generations. Although selected data for several crops are presented for comparison (Table 2), detailed and updated information on these and other species resides in genome databases housed in the U.S. National Agricultural Library, Beltsville, Md. Maps in many species are moderately saturated and incorporate isozymes, RFLPs, and RAPDs. There are at least three maps for potato (*Solanum tuberosum* L.) and rice (*Oryza sativa* L.), and two for bean

and pepper (*Capsicum annuum* L.) that have been constructed using various parents analyzed in diverse generations. Likewise, multiple maps are being developed for other plant species [e.g., *Arabidopsis thalania* (L.) Heynh., corn (*Zea mays* L.), and tomatoes (*Lycopersicon esculentum* Mill.)].

If two or more genetic maps possess a minimal number of common markers they can be merged to create a more informative map (Hauge et al., 1993). However, the type of information (e.g., F_2 vs. BC) and precision of estimates of recombination frequencies (family size) often vary greatly between populations and data sets. Therefore, any procedure that attempts to merge mapping information must "weigh" these types of information to create the "optimal," "most likely" map with the least amount of "internal tension."

A computer program, JoinMap, has recently been developed that considers the estimates of recombination frequency between a given pair of markers of different origin (data sets/mapping populations), calculates and applies the appropriate weighting, and then generates a single recombination value (Stam, 1993). After assigning weights to all available pairwise combinations, JoinMap institutes a numerical search for the best-fitting linear arrangement of the marker loci. JoinMap calculates a goodness-of-fit criterion corresponding to the two hypothesized levels of interference (positive and negative), allowing for an examination of each synthesized map.

Identification of QTLs

In contrast to classical linkage detection for single gene traits, different strategies have been suggested for the identification (i.e., detection and localization) of single QTLs (Edwards et al., 1987; Jiang and Zeng, 1995; Lander and Botstein, 1989). Such strategies attempt to identify major levels of the total genetic variance that contribute to a trait's variation. They differ in approach in the number of markers that they evaluate during linkage estimation. Tests for QTL/trait association can involve the evaluation of one marker at a time, two marker loci simultaneously, or the consideration of all possible marker loci at once. Typical of a one-marker comparison strategy is the use of the one-way analysis of variance (F test) for the analysis of BC progeny, and marker genotype means comparisons (*t* test) for BC and F_2 populations (Soller et al., 1976; Stuber et al., 1992). This approach ignores the potential recombination between a marker and a QTL, and thus will lead to an underestimation of QTL effects if the marker and QTL are not coincident (Edwards et al., 1987). A single marker approach may also incorporate a trait-based analysis in which individuals in the tails of a population distribution are sampled for marker frequencies (Lebowitz et al., 1987). In this case, those markers lying between the tails of the distribution and differing in frequency are assumed to be associated with the QTL affecting the trait.

Approaches which examine two marker loci at once incorporate interval mapping strategies using maximum likelihood for the analysis of single QTLs flanked by a pair of marker loci (Lander and Botstein, 1989; Paterson et al., 1991). The interval approach was developed to take advantage of additional information provided by linkage maps having a relatively high degree of genome saturation (i.e., spacing of markers every 5–20 cM) such as tomato and maize (Paterson et al., 1991; Doebly and Stec, 1991). The interval approach allows for the estimation of putative QTL effects at any location within a marker interval based on the means and variances observed in the marker classes and the recombination frequency between the markers bracketing a particular interval (Lander and Botstein, 1989). This approach is partially limited by its inability to test unlinked markers, and to accurately locate QTLs beyond the terminal markers of a given linkage group.

The consideration of all possible marker loci at once during QTL analysis is complex and involves the regression of trait expression on multiple marker locus values (Cowen, 1989; Rodolphe and Lefort, 1993; Stam, 1993). More recently, interval mapping and multiple regression have been integrated ("hybrid" approaches) to more accurately describe QTL/trait associations (Haley and Knott, 1992; Jansen, 1992, 1993; Jansen and Stam, 1994; Knapp, 1991; Knapp et al., 1990; Martinez and Curnow, 1992; Moreno-Gonzales, 1992; Zeng, 1993, 1994). Regardless of the mapping approach used, the success of MAS depends on the ability to detect QTLs and the consistency of QTLs over environments and generations (Dijkhuizen, 1994; Lande and Thompson, 1990; Shoemaker et al., 1994).

APPLICATION OF MARKERS

Genetic markers have been used effectively in genetic diversity analysis and germplasm organization [e.g., *Arachis* (Lanham et al., 1992), *Brassica* (dos Santos et al., 1994; Thormann et al., 1994), *Vaccinium* (Novy et al., 1994)]; in genetic similarity estimation as predictors of hybrid performance (Bernardo, 1994; Melchinger et al., 1990; Smith et al., 1990); in genetic map construction for the localization of loci conditioning simply inherited traits [e.g., *Pto* locus for resistance to *Pseudomonas syringae* pathovar tomato (*Pst*) (Carland and Staskawicz, 1993); *H1* gene for resistance to *Globodera rostochiensis* (Woll.) Behrens in potato (Gebhardt et al., 1993); *er-1* for resistance to powdery mildew in peas (*Pisum sativum* L.) (Timmerman et al., 1994); downy mildew resistance in lettuce (*Latuca sativa* L.) (Paran and Michelmore, 1993); photoperiod-sensitivity gene in rice (Mackill et al., 1993)]; and QTL analysis (Edwards et al., 1987; Table 3). Marker systems also provide the potential for map-based cloning of specific genes (Tanksley et al., 1995). Although theoretical appraisals of MAS have shown that it could be useful in plant improvement, the application of MAS has not been rigorously evaluated in many crop species (Lande and Thompson, 1990). To date, no cultivar developed through MAS has been publicly released.

Theoretical considerations and computer simulation

Theoretical investigations that probe the potential of MAS are of academic and practical importance. Although there are three general kinds of selection (stabilizing, directional, and disruptive) that could be used by plant breeders, directional selection is preferred because selected phenotypes are distinct from the initial population for economically important attributes. Truncation selection is the simplest type of directional selection. During truncation selection, a phenotypic value is identified as the lower selection limit (truncation point) and individuals are recovered whose phenotypic values are equal to or beyond this value. A prediction equation for response to truncation selection can be defined in terms of response to selection (R; difference in mean phenotype between the progeny generation and the previous generation), heritability (h^2), and the selection differential (S; difference in mean phenotype between the selected parents and the initial population mean) as: $R = h^2S$. Thus, realized heritability can be estimated in the first generation of purely phenotypic selection as: $h^2 = R/S$. Selection intensity (i) or selection differential is often expressed as units of standard deviations (s) in phenotypic value such that $i = S/s$.

Computer-based simulation can allow for tentative interpretation of relatively complex genetic comparisons that have not been previously possible (Edwards and Page, 1994; Lande and Thompson, 1990). Using simple relationships (e.g., $R = h^2S$) and theoretical assumptions of variance components in an initial population, Lande and Thompson (1990) proposed a computer-based simulation model for MAS to estimate genetic effects and gain from index selection. The model provides theoretical estimates for response from truncation selection for QTLs in an F_2 population using a 100-marker loci. The model derives selection indices that maximize the rate of improvement in quantitative characters under various methods of MAS. The model takes into account epistasis by combining multiplicative (multivariant) and classical additive approximations of gene action. Selection is based on an index that incorporates phenotypic and molecular information. The model uses the linkage disequilibrium between molecular marker loci and quantitative trait loci (QTLs) in populations created by a cross between two inbred lines.

Various strategies for plant improvement were tested by Lande and Thompson (1990) using computer simulations to characterize MAS and to provide expectations for phenotypic selection. Potential increases in breeding efficiency through MAS and the population size needed to attain such increases depends on the genetic parameters (i.e., heritability, the proportion of the additive genetic variance explained by the marker loci) and the selection method used. Gain from selection (ΔG) of quantitative traits based on estimated additive effects could be greater for MAS than for phenotypic selection. The relative worth of

Table 3. Estimated number of quantitative trait loci (QTLs) affecting the expression of traits in several crops.

Crop	Population	Trait	QTL (no.)	Range (%) of explained phenotypic σ^{2z}	LOD (range)	Total phenotypic σ^2 explained (%)	Reference
Common bean (*Phaseolus vulgaris* L.)	F_2:F_3	Nodule number	4	1.5–2.8	11.0–17.0	50.0	Nodari et al., 1993b
		Resistance to common blight	4	2.1–6.0	3.5–9.2	75.0	---
Corn (*Zea mays* L.)	F_3BC_1	Grain yield	8	5.6–14.4	6.33–10.86	---	Stuber et al., 1992
	F_3BC_2	Grain yield	6	6.2–18.0	3.16–9.73	---	---
	F_2	European corn borer resistance	7	2.3–9.1	3.4–15.7	38.0	Schon et al., 1993
		Plant height	3	5.7–12.9	10.3–34.1	63.0	---
	F_2:F_3[y]	Ear height	5	6.3–27.8	2.1–5.3	61.2	Veldboom et al., 1994
		Plant height	5	6.4–39.5	2.2–8.2	67.1	---
		GDD to anthesis[x]	6	2.1–5.9	6.2–33.6	62.9	---
		GDD to silk delay	2	2.8–3.2	15.8–17.5	30.9	---
		GDD to silk emergence	5	2.4–11.3	7.8–53.1	80.9	---
	F_2[w]	Number of cupules in single rank	6	4.1–24.6	2.6–11.0	---	Doebley and Stec, 1991
		Tendency of ear to shatter	6	4.3–41.7	2.4–18.6	---	---
		Hardness of outer glume	2	17.5–62.4	7.6–40.6	---	---
		Average length of internodes[v]	5	4.7–45.3	3.0–11.7	---	---
		Number of branches[u]	4	4.3–24.3	2.8–7.6	---	---
		Percent cupules lacking spikelet[t]	5	8.0–25.1	2.9–9.9	---	---
		Number of ears on lateral branch	7	6.3–24.5	2.8–12.9	---	---
		Percent male spikelets[t]	5	5.0–22.5	3.0–15.9	---	---
		Number of rows of cupules	6	5.0–36.0	2.8–15.9	---	---
Cowpea (*Vigna unguiculata* L.)	F_2	Seed weight	4	---	32.0–37	53.0	Fatokun et al., 1992
Potato (*Solanum tuberosum* L.)	BC	Type A trichome browning reaction	2	20.2–52.0	6.5–22.1	63.4	Bonierbale et al., 1994
		Type A trichome density	1	32.0	11.1	32.0	---
		Type A trichome polyphenol oxidase conc.	2	13.2–23.1	2.9–6.3	27.0	---
		Type B trichome sucrose ester levels	5	6.1–49.4	2.0–19.2	67.6	---
		Type B trichome density	2	8.6–35.4	2.9–14.2	38.1	---
	F_1	Late blight	19	---	---	---	Leonards-Schippers et al., 1994
	F_1	Tuber shape	1[s]	---	---	75.0[r]	van Eck et al., 1994
	F_1	Chip color	6	---	5.1–14.9	51.0	Douches and Freyre, 1994
Tomato (*Lycopersicon* spp.)	BC	Fruit mass	6	---	---	58.0	Paterson et al., 1988
		Soluble solids	4	---	---	44.0	---
		Fruit pH	5	---	---	48.0	---
	F_2	Insect resistance	3	---	---	---	Nienhuis et al., 1987
	F_2	Days to first true leaf	3	2.6–13.1	2.4–11.9	18.0	deVicente and Tanksley, 1993
		Days to first flower	7	3.5–10.2	2.2–8.0	43.0	---
		Plant height	9	3.1–8.4	2.2–8.2	42.0	---
		Total number of flower buds	10	2.8–34.0	2.2–34.1	61.0	---
		Number of internodes on primary stem	5	4.8–7.4	3.2–5.71	23.0	---
		Total number of internodes	9	3.7–12.8	2.8–10.8	52.0	---
		Number of well developed branches	8	3.1–9.5	2.7–7.5	53.0	---
		Total plant fresh weight	2	3.0–4.4	2.2–2.74	7.0	---
		Total plant dry weight	5	3.2–7.0	2.3–3.5	21.0	---
	F_2:F_3	Soluble solids	7	3.0–12.0	6.0–28.0	44.0	Paterson et al., 1991
		Fruit mass	11	2.3–21.5	4.0–42.0	72.0	---
		Fruit pH	9	2.4–6.1	4.2–28.0	34.0	---

[z]Given on a per locus basis.
[y]F_2 classified by F_3 families.
[x]Growing degree days.
[w]*Zea mays*; corn x teosinte.
[v]In primary lateral branches.
[u]In primary lateral inflorescence.
[t]Pedicellate spikelet.
[s]Multiple alleles were detected.
[r]Total genetic variance.

MAS is greatest for characters with low heritability when additive genetic variance is associated with the marker. More recently, Gimelfarb and Lande (1994b) have demonstrated that this same logic could be applied to nonadditive characters.

The simplifying assumptions of computer-based models, however, can lead to over- or under-estimations of R. For instance, fitness plays an integral part in response of individuals to selection. Heterotic advantage can be defined in terms of fitness. In many species, heterosis is pivotal to the expression of average fitness in a population. Estimates of R are only valid if an individual's "fitness" is interpreted as the probability that an individual is included among the group selected as parents in the next generation (Gimelfarb, 1989).

The potential utility of MAS in practical plant breeding programs is limited by: 1) the number of molecular marker loci required to detect all significant linkage associations; 2) population sizes required to detect QTLs for traits with low heritability; 3) the sampling errors associated with the weighting of indices when combining molecular marker loci; and 4) phenotypic information and the cost per

unit information gained (Edwards and Page, 1994; Lande and Thompson, 1990). Marker effectiveness (i.e., selection response) increases as the number of QTLs affecting a trait decreases (Edwards and Page, 1994). The effectiveness of MAS decreases as the linkage distance between markers and QTLs increases. Moreover, greater genetic gain can be made when flanking QTLs between two marker loci are used as compared to single markers, if single markers are loosely linked to a QTL. However, the use of flanking markers requires the characterization of twice as many markers as compared to selection using single markers. Thus, where dense maps are available, the value (cost/unit information) of flanking markers decreases as marker QTL associations increase.

The effectiveness of MAS is also determined by the relative linkage disequilibria between the marker loci and QTLs that condition trait expression (Lande and Thompson, 1990). Linkage disequilibria (between genetic markers and QTLs) is maximized by the mating germplasm of divergent origin. Such matings occur regularly during plant improvement (e.g., crossing between elite inbreds to begin genetic recombination for line development). Fixation of desirable trait loci in an elite background is the goal of breeding programs. Greater genetic gain is likely when fewer genes are involved in trait behavior because less recombination occurs (Edwards and Page, 1994). Having many QTLs exacerbates the problem of marker-QTL recombination, and thus the time required for fixation increases as the number of QTLs associated with a particular trait increases.

The multiple regression of phenotype on genetic markers can be used during MAS to provide a tactical assessment of gain from selection. Analysis of such relationships during MAS capitalizes on the linkage disequilibrium generated by the original mating of two inbred lines (Gimelfarb and Lande, 1994a). MAS can, therefore, be very effective during the early generations of population improvement where important linkages have not been eroded by recombination (Edwards and Page 1994; Lande and Thompson, 1990). Markers that contribute significantly to selection in initial generations should be re-evaluated each generation to determine their continued effectiveness (Gimelfarb and Lande, 1994a). This need can be costly and undermine the potential usefulness of MAS.

If a significant amount of the additive variance associated with a QTL can be accounted for by molecular markers, then MAS can increase breeding efficiency (Edwards and Page, 1994; Gimelfarb and Lande, 1994b). Likewise, the effectiveness of marker loci will be increased as the number of individuals in a population is increased, since a greater proportion of the additive genetic variance can potentially be explained. When trait heritability is low, population size must be relatively large (100–1000 individuals) to include unrelated individuals that detect additive variance associated with marker loci. Moreover, full- or half-sib populations are likely to require larger samples, depending on the degree of dominance associated with the trait.

Sample size is important when considering the potential loss of efficiency in MAS during protracted index selection (Gimelfarb and Lande, 1994a). Sampling errors can occur during model building as the relative weights of molecular and phenotypic information are estimated. MAS may only be cost efficient if phenotypic selection is made difficult due to large environmental effects and/or the number of loci affecting such traits is large (Lande and Thompson, 1990). However, increasing the number of markers that contribute to the selection index does not necessarily increase the effectiveness of MAS. The use of many markers may in fact result in a weaker response to selection.

Application

Simmonds (1979) has stated that "...plant breeding often not only generates benefits but is also attractive in having relatively low implementation costs..." This statement may no longer apply if MAS is rigorously applied to crop improvement. The costs of MAS can be high when compared to classical phenotypic selection, and the cost : benefit ratio may not be high enough to warrant use of MAS. The cost-benefit relationship can now be more critically evaluated in differing marker systems (Table 1). For instance, Lande (1992) indicated that the cost of scoring RFLPs is on the order of 100 to 1000 times as expensive as measuring standard phenotypes in most crops. Simulation experiments by Ragot and Hoisington (1993) indicated that the costs for employing RAPDs may be higher than for RFLPs as the number of individuals and markers in an experiment are increased. While RAPDs were found to be most cost/time efficient when sample sizes were small, RFLPs were more advantageous as sample sizes were increased. Darvasi and Soller (1994) considered experimental costs (i.e., number of individuals and marker spacing) that might be incurred during QTL analysis of a plant species possessing a genome size of 1000 cM. They concluded that the costs of MAS would be prohibitive since hundreds of individuals would be needed in a typical marker-QTL experiment, even if all possible markers could be used and the power of QTL detection (i.e., LOD) was high.

Although the costs associated with MAS are currently high, it has been shown to have potential utility for managing complex traits (Table 3). For instance, Stuber and Edwards (1986) recorded similar genetic gains when using either phenotypic selection or MAS (isozymes) for quantitative traits in maize. Similar results were found by Stromberg (as cited by Dudley, 1993). More recently, Stuber (1995) observed that significant genetic gain could be obtained during marker-assisted backcrossing in maize aimed at transferring targeted QTLs from T×303 into B73 and from Oh43 into Mo17. Although the yield of hybrids between the "enhanced" B73 and Mo17 exceeded that of control hybrids by more than 15%, no parallel assessment of the relative efficiency of classical breeding was included. Edwards and Johnson (1994) used MAS in two sweet corn populations (A and B) and found a positive response from selection (two cycles per year) for several traits. Elite lines were crossed to produce F_4 lines that were then subjected to either MAS (RFLPs) or phenotypic recurrent selection (PRS) for yield and quality traits. These original parents and the resultant populations (MAS and PSR) were crossed to two testers for replicated comparison at one location in one year. Although positive response to selection was observed for six of 11 traits in one population (A), differences between the average performances of the hybrids developed from MAS lines and parental hybrids were not significant. Likewise, the overall response to phenotypic or MAS selection was similar in a second population (B). The authors hypothesized that this lack of response to MAS was due to antagonistic effects of genome regions responsible for the yield and quality traits observed.

SUMMARY

Molecular markers and associated technologies can assist in map construction and the analysis of the molecular and genetic basis of quantitative and qualitative traits. Molecular markers that are tightly linked to economically important traits that are under the control of single genes have potential for immediate utility in plant improvement programs. Perhaps the most optimistic prospects for MAS is in disease resistance breeding, especially where several genes control resistance with complex interactions and where pyramiding of genes is desirable (Kelly, 1995; Schafer and Roelfs, 1985; Stavely et al., 1989).

Optimal use of QTL regions will, however, require a knowledge of their often complex epistatic interactions. Moreover, although MAS may have potential in population and inbred line development, it will likely have little or no effect in reducing the need for replicated field trials and testing. Optimal methods for mapping QTLs are still being debated and more sophisticated computer-aided analysis procedures are being developed. When QTLs and single genes are adequately mapped, they can be isolated biochemically (Tanskley et al., 1995; Young, 1990). Methods for their isolation (i.e., cloning) and characterization are also points of considerable discussion.

The effectiveness of any MAS procedure will depend on the accuracy of the phenotypic classification of trait expression and the degree of linkage between a marker(s) and traits of interest. Although MAS may increase gain from selection when compared to phenotypic selection, marker utility in plant improvement programs will ultimately be determined by cost/unit information (Edwards and Page, 1994). Clearly, laboratory costs associated with MAS applications are decreasing, and more effective and efficient molecular markers are being developed (Gu et al., 1995). This progress will make MAS more attractive and

will foster its prudent implementation as a tool for plant breeding. As a result of such changes, MAS might have potential for selection of characters such as yield components in agronomic and horticultural crops. Nevertheless, in horticultural crops, where many complex and highly integrated aesthetic, culinary, and organoleptic attributes are considered necessary for market acceptance, plant breeding expertise and decision making ability will clearly remain pivotal for genetic improvement.

Literature Cited

Akkaya, M.S., A.A. Bhagwat, and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139.

Albini, S.M. and G.H. Jones. 1987. Synoptonemal complex spreading in *Allium cepa* and *A. fistulosum*. I. The initiation and sequence pairing. Chromosoma 95:324–338.

Allard, R.W. 1956. Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278.

Beckmann, J.S. and M. Soller. 1983. Restriction fragment length polymorphisms in genetic improvement: Methodologies, mapping and costs. Theor. Appl. Genet. 67:35–43.

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

Bonierbale, M.W., R.L. Plaisted, O. Pineda, and S.D Tanksley. 1994. QTL analysis of trichome-mediated insect resistance in potato. Theor. Appl. Genet. 87:973–987.

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

Caetano-Anolles, D., B.J. Bassam, and P.M. Gresshoff. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Biotechnology 9:553–557.

Carland, F.M. and B.J. Staskawicz. 1993. Genetic characterization of the *Pto* locus of tomato: Semi-dominance and cosegregation of resistance to *Pseudomonas syringae* pathovar tomato and sensitivity to the insecticide Fenthion. Mol. Gen. Genet. 239:17–27.

Causse, M.A., T.M. Fulton, Y.G. Cho, S.N. Ahn, J. Chunwongse, K. Wu, J. Xiao, Z. Yu, P.C. Ronald, S.E. Harrington, G. Second, S.R. McCouch, and S. Tanksley. 1994. Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251–1274.

Chakravarti, A., L.K. Lasher, and J.E. Reffer. 1991. A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128:175–182.

Churchill, G.A. and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Cowen , N.M. 1989. Multiple regression analysis of RFLP data sets used in mapping QTLs, p. 113–116. In: T. Heletjaris and B. Burr (eds.). Development and applications of molecular markers to problems in plant genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Crow, J.F. 1990. Mapping functions. Genetics 125:669–671.

Darvasi, A. and M. Soller. 1994. Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. Theor. Appl. Genet. 89:351–357.

de Vincente, M.C. and S.D. Tanksley. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596.

Diers, B.W., P. Klein, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 83:608–612.

Dijkhuizen, A. 1994. Application of restriction fragment length polymorphism for the assessment of genetic variability and study of quantitatively inherited traits in cucumber (*Cucumis sativus* L.). PhD Diss., Univ. of Wisconsin, Madison.

Doebly, J. and A. Stec. 1991. Genetic analysis of morphological differences between maize and teosinte. Genetics 129:285–295.

dos Santos, J.B., J. Nienhuis, P. Skroch, J. Tivang, and M.K. Slocum. 1994. Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. Theor. Appl. Genet. 87:909–915.

Douches, D.S. and R. Freyre. 1994. Identification of genetic factors influencing chip color in diploid potato (Solanum spp.) Amer. Potato J. 71:581–590.

Dudley, J.W. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Sci. 33:660–668.

Echt, C., S. Knapp, and B.H. Liu. 1992. Genome mapping with non-inbred crosses using GMendel 2.0. Maize Genet. Coop. Nwsl. 66:27–29.

Edwards, M. and L. Johnson. 1994. RFLPs for rapid recurrent selection. Proc. Symp. Analysis of Molecular Marker Data, p. 33–40. Aug. 1994. Corvallis, Ore.

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

Edwards, M.D., C.W. Stuber, and J.F. Wendel. 1987. Molecular facilitated investigations of quantitative trait loci in maize. I. Number, genomic distribution and types of gene action. Genetics 116:113–125.

Fatokun, C.A., D.I. Menancio-Havtea, D. Danesh, and N.D. Young. 1992. Evidence for orthologus seed weight genes in cowpea and mung bean based on RFLP mapping. Genetics 132:841–846.

Faure, S., J.L. Noyer, J.P. Horry, F. Bakry, C. Lanaud, and D. Gonzales de Leon. 1993. A molecular marker-based linkage map of diploid banana. Theor. Appl. Genet. 87:517–526.

Ferreira, M.E., P.H. Williams, and T.C. Osborn. 1994. RFLP mapping of *Brassica napus* using double haploid lines. Theor. Appl. Genet. 89:615–621.

Fisher, R.A. 1946. A system of scoring linkage data with special reference to pied factors in mice. Amer. Naturalist 80:568–578.

Fregeau, C.J. and R.M. Fourney. 1993. DNA typing with fluorescently tagged short tandem repeats: A sensitive and accurate approach to human identification. BioTechniques 15:100–119.

Gebhardt, C., D. Mugnieri, E. Reitter, F. Salamini, and E. Bonnel. 1993. Identification of RFLP markers closely linked to the H1 gene to conferring resistance to *Globodera rostochiensis* in potato. Theor. Appl. Genet. 85:541–544.

Gimelfarb, A. 1989. Genotypic variation maintained under stabilizing selection without mutations: Epistasis. Genetics 123:217–227.

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

Gimelfarb, A. and R. Lande. 1994b. Simulation of marker assisted selection for non-additive traits. Genetical Res. 64:127–136.

Gu, W.K., N.F. Weeden, J. Yu, and D.H. Wallace. 1995. Large-scale, cost-effective screening of PCR products in marker-assisted selection applications. Theor. Appl. Genet. 91:465–470.

Gupta, M., Y.S. Chyi, J. Romero-Severson, and J.L. Owen. 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theor. Appl. Genet. 89:998–1006.

Haldane, J.B.S. 1919. The combination of linkage values, and the calculation of distance between the loci of linked factors. J. Genet. 8:299–309.

Haldane, J.B.S. and C.A.B. Smith. 1947. A new estimate of the linkage between the genes for colour-blindness and haemophilia in man. Ann. Eugenics 14:10–31.

Haley, C.S. and S.A. Knot. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324.

Haley, S.D., L.K. Afanador, and J.D. Kelly. 1994. Selection for monogenic resistance traits with coupling- and repulsion-phase RAPD markers. Crop Sci. 34:1061–1066.

Hauge, B.M., S.M. Hanley, S. Cartinhour, J.M. Cherry, H.M. Goodman, M. Koorneef, P. Stam, C. Chang, S. Kempin, L. Medrano, and E.M. Meyerowitz. 1993. An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J. 3:745–754.

Jansen, R.C. 1992. A general mixture model for mapping quantitatitve trait loci by using molecular markers. Theor. Appl. Genet. 85:252–260.

Jansen, R.C. 1993. Interval mapping of multiple quantitative trait loci. Genetics 135:205–211.

Jansen, R.C. and P. Stam. 1994. High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455.

Jarrell, D.C., M.L. Roose, S.N. Traugh, and R.S. Kupper. 1992. A genetic map of citrus based on the segregation of isozymes and RFLPs in an intergeneric cross. Theor. Appl. Genet. 84:49–56.

Jarvis, P., C. Lister, V. Szabo, and C. Dean. 1994. Integration of CAPs markers into the RFLP map generated using recombinant inbred lines of *Arabidopsis thaliana*. Plant Mol. Biol. 24:685–687.

Jiang, C. and Z. Zeng. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111–1127.

Kelly, J.D. 1995. Use of random amplified polymorphic DNA markers in breeding from major gene resistance to plant pathogens. HortScience 30:461–465.

Kennard, W.C., K. Poetter, A. Dijkhuizen, V. Meglic, J. Staub, and M. Havey. 1994. Linkages among RFLP, RAPD, isozyme, disease resistance, and morphological markers in narrow and wide crosses of cucumber. Theor. Appl. Genet. 89:42–48.

Kleinhofs, A., A. Kilian, M.A. Shagai Maroof, R.M. Biyashev, P. Hayes, F.Q. Chen, N. Lapitan, A. Fenwick, T.K. Blake, V. Kanazin, E. Ananiev, L. Dahleen, D. Kudrna, J. Bollinger, S.J. Knapp, B. Liu, M. Sorrels, M. Heun, J.D. Franckowiak, D. Hoffman, R. Skadsen, and B.J. Steffenson. 1993. A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor. Appl. Genet. 86:705–712.

Knapp, S.J. 1991. Using molecular markers to map multiple quantitative trait loci: Models for backcross, recombinant inbred, and doubled haploid progeny. Theor. Appl. Genet. 81:333–338.

Knapp, S.J., W.C. Bridges, and D. Birkes. 1990. Mapping quantitative trait loci using molecular marker linkage maps. Theor. Appl. Genet. 79:583–592.

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

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugenics 12:172–175.

Kramer, H.H. and C.R. Burnham. 1947. Methods of combining linkage intensity values from backcross, F_2 and F_3 genetic data. Genetics 32:379–390.

Lande, R. 1992. Marker-assisted selection in relation to traditional methods of plant breeding, p. 437–451. In: H.T. Stalker and J.P. Murphy (eds.). Plant breeding in the 1990s. CAB Intl., Wallingford, U.K.

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

Lander, E.S. and D. Botstein. 1986. Mapping complex genetic traits in humans: New method using a complex RFLP linkage map. Cold Spring Harbor Symposia on Quantitative Biology. 51:49–62.

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

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln, and L. Newberg. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

Landry, B.S., N. Hubert, T. Etoh, J.J. Harada, and S.E. Lincoln. 1991. A genetic map for *Brassica napus* based on restriction fragment length polymorphism detected with expressed DNA sequences. Genome 34:543–552.

Lanham, P.G., S. Fennell, J.P. Moss, and W. Powell. 1992. Detection of polymorphic loci in *Arachis* germplasm using random amplified polymorphic DNAs. Genome 35:885–889.

Lark, K.G., K. Chase, F. Adler, L.M. Mansur, and J.H. Orf. 1995. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele. Proc. Natl. Acad. Sci. USA 92:4656–4660.

Lebowitz, R.J., M. Soller, and J.S. Beckmann. 1987. Trait-based analysis for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. Theor. Appl. Genet. 73:556–562.

Leonards-Schippers, C., W. Gieffers, R. Schäfer-Pregl, E. Ritter, S.J. Knapp, F. Salamini, and C. Gebhardt. 1994. Quantitative resistance to *Phytophthora infestans* in potato: A case study for QTL mapping in an allogamous plant species. Genetics 137:67–77.

Mackill, D.J., M.A. Salam, Z.Y. Wang, and S.D. Tanksley. 1993. A major photoperiod-sensitivity gene tagged with RFLP and isozyme markers in rice. Theor. Appl. Genet. 85:536–540.

Manly, K.F. and R.W. Elliot. 1991. MapManager, a microcomputer program for analysis of data from recombinant inbred strains. Mammalian Genome 1:123–126.

Markert, C.L. and F. Moller. 1959. Multiple forms of enzymes: Tissues, ontogenetic, and species specific patterns. Proc. Natl. Acad. Sci. USA 45:753–763.

Martinez, O. and R.N. Curnow. 1992. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. 85:480–488.

Mather, K. 1938. Measurement of linkage in heredity. Methuen and Co., London.

Melchinger, A.E., M. Lee, K.R. Lamkey, A.R. Hallauer, and W.L. Woodmann. 1990. Genetic diversity for restriction fragment length polymorphism and heterosis for two diallel sets of maize inbreds. Theor. Appl. Genet. 80:488–496.

Michelmore, R.W., I. Paran, and R.V. Kesseli. 1991. Identification of markers linked to disease resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions using segregating populations. Proc. Natl. Acad. Sci. USA 88:9828–9832.

Moreno-Gonzales, J. 1992. Genetic models to estimate additive and non-additive effects of marker-associated QTL using multiple regression techniques. Theor. Appl. Genet. 85:435–444.

Morgante, M., A. Rafalski, P. Biddle, S. Tingey, and A.M. Olivieri. 1994. Genetic mapping and variability of seven soybean simple sequence repeat loci. Genome 37:763–769.

Nienhuis, J., T. Helentjaris, M. Slocum, B. Ruzzero, and M. Schaefer. 1987. Restriction fragment length polymporphism analysis of loci associated with insect resistance in tomato. Crop Sci. 27:797–803.

Nodari, R.O., S.M. Tscai, R.L. Gilberston, and P. Gepts. 1993a. Towards an integrated linkage map of common bean. 2. Development of an RFLP-based linkage map. Theor. Appl. Genet. 85:513–520.

Nodari, R.O., S.M. Tsai, P. Guzman, R.L. Gilbertson, and P. Gepts. 1993b. Toward an integrated linkage map of common bean. 3. Mapping genetic factors controlling host–bacteria interaction. Genetics 134:341–350.

Nordheim, E.V., D.M. O'Malley, and S.C. Chow. 1984. On the performance of a likelihood ratio test for genetic linkage. Biometrics 40:785–790.

Novy, R.G., C. Kobak, J. Goffreda, and N. Vorsa. 1994. RAPDs identify varietal misclassification and regional divergence in cranberry [*Vaccinium macrocarpon* (Ait.) Pursh]. Theor. Appl. Genet. 88:1004–1010.

Olson, M., L. Hood, C. Cantor, and D. Botstein. 1989. A common language for physical mapping of the human genome. Science 245:1434–1435.

Owen, J.L. and C.M. Uyeda. 1991. Single primer amplification of avian genomic DNA detects polymorphic loci. Ann. Biotechnol. 2:107–122.

Paran, I., R. Kesseli, and R. Michelmore. 1991. Identification of RFLP and RAPD markers linked to downy mildew resistance genes in lettuce using near-isogenic lines. Genome 34:1021–1027.

Paran, I. and R.W. Michelmore. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor. Appl. Genet. 85:985–993.

Paterson, A.H., S. Damon, J.D. Hewitt, D. Zamir, H.D. Rabinowitch, S.E. Lincoln, E.S. Lander, and S.D. Tanksley. 1991. Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations, and environments. Genetics 127:187–197.

Paterson, A.H., E.S. Lander, J.D. Hewitt, S. Peterson, S.E. Lincoln, and S.D. Tanksley. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726.

Philipp, U., P. Wehling, and S.G. Wricke. 1994. A linkage map of rye. Theor. Appl. Genet. 88:243–248.

Pillen, K., G. Streinruchen, G. Wricke, R.G. Herrmann, and C. Jung. 1992. A linkage map of sugar beet (*Beta vulgaris* L.). Theor. Appl. Genet. 84:129–135.

Rafalski, J.A. and S.V. Tingey. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet. 9:275–279.

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

Reinisch, A.J., J. Dong, C.L. Brubaker, D.M. Stelly, J.F. Wendel, and A.H. Paterson. 1994. A detailed RFLP map of cotton, *Gossypium hirsutum* x *Gossypium barbadense*: Chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829–847.

Reiter, R.S., J.G.K. Williams, K.A. Feldman, A. Rafalski, S.V. Tingey, and P.A. Scolnik. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplied polymorphic DNAs. Proc. Natl. Acad. Sci. USA 89:1477–1481.

Risch, N. 1992. Genetic linkage: Interpreting LOD scores. Science 255:804.

Rodolphe, F. and M. Lefort. 1993. A multi-marker model for detecting chromosomal segments displaying QTL activity. Genetics 134:1277–1288.

Schafer, F.R. and A.P. Roelfs. 1985. Estimated relation between numbers of urediniospores of *Puccina graminis* f. sp. *tritici* and rates of occurrence of virulence. Phytopathology 75:749–750.

Schon, C.C., M. Lee, A.E. Melchinger, W.D. Guthrie, and W.L. Woodman. 1993. Mapping and characterization of quantitative trait loci affecting resistance against second generation European corn borer in maize with the aid of RFLPs. Heredity 70:648–659.

Shoemaker, R.C., L.L. Lorenzen, B.W. Diers, and T.C. Olson. 1994. Genome mapping and agriculture, p. 1–10. In: P.M. Gresshoff (ed.). Plant genome analysis. CRC Press, Boca Raton, Fla.

Shute, N.C.E. 1988. Power properties of the likelihood ratio test in an analysis of genetic diseases. Biometrics 44:951–958.

Simmonds, N.W. 1979. Principles of crop improvement. Longman Group, London.

Smith, O.S., J.S.C. Smith, S.L. Bowen, R.A. Tenborg, and S.J. Wall. 1990. Similarities among a group of elite maize inbreds as measured by pedigree. F_1 grain yield, heterosis and RFLPs. Theor. Appl. Genet. 80:833–840.

Soller, M., T. Brody, and A. Genizi. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47:35–39.

Song, K.M., J.Y. Suzuki, M.K. Slocum, P.H. Williams, and T.C. Osborn. 1991. A linkage map of *Brassica rapa* (syn. *compestris*) based on restriction fragment length polymporphism loci. Theor. Appl. Genet. 82:296–304.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.

Stam, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J. 3:739–744.

Stansfield, W.D. 1969. Theory and problems of genetics. Schaum's outline series. McGraw-Hill, New York.

Staub, J.E and L. Crubaugh. 1995. Selection for multiple lateral determinate cucumber genotypes. Cucurbit Gen. Coop. Rpt. 18:5–6.

Staub, J.E., L.J. Kuhns, B. May, and P. Grun. 1982. Stability of potato tuber isozymes under different storage regimes. J. Amer. Soc. Hort. Sci. 107:405–408.

Stavely, J.R., J.R. Steadman, and R.J. McMillian, Jr. 1989. New pathogenic variabiliy in *Uromyces appendiculatus* in North America. Plant Dis. 73:428–432.

Stuber, C. 1995. Yield improvement success using marker-facilitated QTL transfer between maize lines, p. 13. In: Plant Genome III. Abstr. Intl. Conf. Status of Plant Genome Res. 15–19 Jan. 1995, San Diego.

Stuber, C. and M.D. Edwards. 1986. Genotypic selection for improvement of quantitative traits in corn using molecular marker loci. Rpt. Ann. Corn Sorghum Res. Conf. Washington, D.C. 41:70–83.

Stuber, C.W., S.E. Lincoln, D.W. Wolf, T. Helentjaris, and E. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839.

Suiter, K.A., J.F. Wendel, and J.S. Case. 1983. Linkage-1: A pascal computer program for the detection and analysis of genetic linkage. J. Hered. 74:203–204.

Swanson, C.P., T. Merz, and W.J. Young. 1990. Cytogentics: The chromosome in division, inheritance and evolution. Prentice Hall, New Delhi, India.

Tanksley, S.D., M.W. Ganal, and G.B. Martin. 1995. Chromosome landing: A paradigm for map-based gene cloning in plants with large genomes. Trends Genet. 11:63–68.

Tanksley, S.D., M.W. Ganal, J.P. Prince, M.C. de Vicente, M.W. Bonierbale, P. Broun, T.M. Fulton, J.J. Giovannoni, S. Grandillo, G.B. Martin, R. Messeguer, J.C. Miller, L. Miller, A.H. Paterson, O. Pineda, M.S. Roder, R.A. Wing, W. Wu, and N.D. Young. 1992. High density molecular maps of the tomato and potato genomes. Genetics 132:1141–1160.

Tanksley, S.D., J. Miller, A. Paterson, and R. Bernatsky. 1988. Molecular mapping of plant chromosomes, p. 157–173. In: J.P. Gustafson and R. Appels (eds.). Chromosome structure and function: Impact of new concepts. Plenum Press, New York.

Teutonico, R.A. and T.C. Osborn. 1994. Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage map of *B. napus*, *B. oleracea* and *Arabidopsis thaliana*. Theor. Appl. Genet. 89:885–894.

Thormann, C.E., M.E. Ferreira, L.E. Camargo, J.G. Tivang, and T.C. Osborn. 1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88:973–980.

Timmerman, G.M., T.J. Frew, N.F. Weeden, A.L. Miller, and D.S. Goulden. 1994. Linkage analysis of *er-1*, a recessive *Pisum sativum* gene for resistance to powdery mildew fungus (*Erysiphe pisi* D.C.). Theor. Appl. Genet. 88:1050–1055.

Vallejos, C.E., N.S. Sakiyama, and C.D. Chase. 1992. Molecular marker-based linkage map in *Phaseolus vulgaris* L. Genetics 131:733–740.

van Eck, H.J., J.M.E. Jacobs, P. Stam, J. Ton, W.J. Stickema, and E. Jacobsen. 1994. Multiple alleles for tuber shape in diploid potato detected by qualitative and quantitative genetic analysis using RFLPs. Genetics 137:303–309.

Veldboom, L.R., M. Lee, and W.L. Woodman. 1994. Molecular marker facilitated studies in an elite maize population: 1. Linkage analysis and determination of QTLs for morphological traits. Theor. Appl. Genet. 88:7–16.

Webb, D.M., S.J. Knapp, and L.A. Tagliani. 1992. Restriction fragment length polymorphism and allozyme linkage map of *Cuphea lanceolata*. Theor. Appl. Genet. 83:528–532.

Weeden, N.F., M. Timmerman, M. Hermmat, B.E. Kneen, and M.A. Lodhi. 1992. Inheritance and reliability of RAPD markers. In: J. Nienhuis (ed.). Proc. Symp. Applications of RAPD Technology to Plant Breeding, p. 12–17. Nov. 1992, Minneapolis, Minn.

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

Whitkus, R., J. Doebley, and M. Lee. 1992. Comparative genome mapping of sorghum and maize. Genetics 132:1119–1130.

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

Xu, G.W., C.W. Magill, K.F. Schertz, and G.E. Hart. 1994. A RFLP linkage map of *sorghum bicolor* (L.) Moendr. Theor. Appl. Genet. 89:139–145.

Young, N. and S.D. Tanksley. 1989. RFLP analysis of the size of chromosomal segments retained around the *Tm-2* locus of tomato during backcross breeding. Theor. Appl. Genet. 77:353–359.

Young, N.D. 1990. Potential applications of map-based cloning to plant pathology. Physiol. Mol. Plant Pathol. 37:81–94.

Zabeau, M. and P. Vos. 1993. Selective restriction fragment amplification: A general method for DNA fingerprints. European Patent Application. Publ. 0534858A1.

Zamir, D. and Y. Tadmor. 1986. Unequal segregation of nuclear genes in plants. Bot. Gaz. 147:355–358.

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.

Zietkiewicz, E., A. Rafalski, and D. Labuda. 1994. Genome fingerprinting by simple sequence repeat (SSR)–anchored polymerase chain-reaction amplification. Genomics 20:176–183.

Appendix: Definitions of abbreviations for genetic terms used.

AFLPs—amplified fragment length polymorphisms. A class of genetic markers produced by selective amplification of restriction enzyme-digested DNA fragments. Primers are designed from the adapters which are ligated to restriction enzyme-digested DNA fragments.

AP-PCR—arbitrary-primed PCR. Amplify discrete patterns by employing single primers of 10 to 50 bases in length in PCR of genomic DNA. The first two cycles of PCR are carried out under nonstringent conditions. The final products are structurally similar to the RAPD products.

ASAPs—allele-specific associated primers. Specific primers that are used in a fluorescence-based procedure to amplify DNA template in microtiter plates to generate only a single DNA fragment (per primer used) at stringent annealing temperatures (similar to SCARs).

BSA—bulked segregant analysis. A genetic analysis procedure that identifies linkage associations by using two bulked DNA samples drawn from a segregating population. These bulks contain individuals that are identical for a particular trait or genomic region but arbitrary at all unlinked regions (i.e., heterozygous), and thus DNA polymorphisms between bulks indicate potential linkage.

CAPs—cleaved amplified polymorphic sequences. Generated by restriction enzyme digestion of PCR products. Such digests are compared for their differential migration during electrophoresis.

CC—coefficient of coincidence. A genetic measure (value) that denotes the proportion of observed double crossovers to the number expected from the random combination of single crossovers among linked genes.

cM—centimorgan. A measure of genetic distance that is used to indicate a separation between two genes using genetic recombination as a base estimator. One unit of distance (cM) is equivalent to 1% crossing over.

DAF—DNA amplification fingerprinting. Produced by employing single arbitrary primers as short as 5 bases in PCR of genomic DNA. The amplified patterns are complex and structurally similar to the RAPD products.

ΔG—genetic gain. A measure of response to selection. The average phenotype *M* reflects the effects of the average genotype ($\bar{G}_1$) and the population mean M´ reflects the average genotypic mean of the offspring, $\bar{G}_2$. The difference $\bar{G}_2 - \bar{G}_1$ is called genetic gain from selection.

DGGE—denaturing gradient gel electrophoresis. A polyacrylamide technique that is used for the identification of small sequence differences created by single base substitutions and/or deletions. Samples (e.g., PCR products such as RAPDs) are loaded onto a denaturing gel where DNA sequences migrate differently based on small base-pair differences.

h^2—heritability. The ratio of the genetic variability to the total variability of a character or a trait in a population. Heritability is a correlation between genotype and phenotype.

i—selection intensity. Selection differential (S) expressed in units of standard deviations (s) in phenotypic value (i = S/s).

LOD—likelihood odds ratio. A measure used to test a null hypothesis that no linkage exists between two genes. Recombination value is used in a maximization expression to determine the likelihood (L) of association of linkage between genetic loci. The odds ratio of maximization event is given as: L(x)/L(0.5). This form can be converted to the log of the odds ratio (LOD) where LOD = log[L(x)/L(0.5)].

MAS—marker-assisted selection. The use of genetic markers linked to traits of interest for phenotypic selection.

mf—genetic mapping function. A formula expressing the quantitative relationship between distance in a linkage map (crossing over; additive) and recombination frequency (nonadditive) such that recombination frequency is transformed into additive metric value relating to the number of crossovers.

NILs—nearly isogenic lines. Highly inbred lines constructed by repeated self-pollination or by backcrossing to a recurrent parent.

PCR—polymerase chain reaction. A technique used for enzymatic in vitro amplification of specific DNA sequences present between two convergent oligonucleotide primers that hybridize to opposite DNA strands.

QTL—quantitative trait loci. Multiple loci or genomic regions that affect trait expression.

r—recombination fraction. A measure of the frequency of exchange or crossing over between two specific loci.

R—response to selection. Generally defined (truncation selection) as $R = h^2S$; where S = selection diffferential and h^2 = heritability.

RAPD—random amplified polymorphic DNA. DNA produced by using single short oligonucleotide primers (usually 9 to 10 nucleotides) in PCR of genomic DNA. A discrete set of amplified DNA represents target sequences bounded by opposite oriented primer annealing sites.

RFLPs—restriction fragment length polymorphisms. Variations in the length of DNA fragments that have been generated by DNA digestion with different resistriction enzymes (i.e., restriction endonuclease). Such variations in the genomic DNA are visualized by hybridization to labeled homologous DNA sequences called probes.

S—selection differential. A measure of the selection intensity. Given that *M* is the initial mean of a population and M_s is the mean of the parents that have been selected for the next generation, $S = M_s - M$.

SCARs—sequenced characterized amplified regions resulting from the sequencing of RAPD marker termini and the designing of longer primers (e.g., 24 nt) for the specific amplification of one band.

SPAR—single primer amplification reaction. A DNA marker system that uses primers based on microsatellites or simple sequence repeats (SSRs) and amplifies inter-SSR DNA sequences.

SSCP—single-strand conformational polymorphism. Identified using polyacrylamide techniques. Single-strand DNAs fold differently due to conformational differences due to sequence variation (single base pair substitutions or deletions) and thus migrate differently.

SSRs—simple sequence repeats. Tandem arrays of 2 to 5 base repeat units (in particular dinucleotide repeats) in plants.

STRs—short tandem repeats. Tandem arrays of 2 to 5 base repeat units (in particular dinucleotide repeats) in vertebrates.