

Does Marker-assisted Selection Make Dollars and Sense in a Fruit Breeding Program?

James J. Luby¹

Department of Horticultural Science, University of Minnesota, St. Paul, MN 55108

Douglas V. Shaw

Department of Pomology, University of California, Davis, CA 95616

Numerous authors have alluded to the potential benefits of marker-assisted selection (MAS) for fruit breeding programs (e.g., Baird et al., 1996; Conner et al., 1997; Dirlewanger et al., 1998; Foolad et al., 1995; Gianfranceschi et al., 1996; Gill et al., 1998; Gmitter et al., 1998; Hemmat et al., 1997; Lahogue et al., 1998; Maliepaard et al., 1998; Markussen et al., 1995; Wang et al., 1998). The cost-benefit relationships are presumed to be more favorable for many fruit crops than for annual crops because of their large size and long juvenile phase. Since the fruit is the item of economic value, many traits cannot be evaluated until seedlings mature. Thus, MAS during the juvenile phase has been proposed to speed the selection process or reduce the progeny sizes and the costs of carrying individuals to maturity in the field.

Despite the perceived benefits of MAS, efforts to develop maps and establish marker-trait linkages in fruit and nut crops lag behind those in many other crops (Paterson, 1996). In several fruit crops, linkages between markers and single genes affecting fruit traits, morphology, or disease or pest resistance have been established (e.g., Arus et al., 1999; Bartish and Weeden, 1999; Conner et al., 1997; Dirlewanger et al., 1998; Gardiner et al., 1999; Gianfranceschi et al., 1996; Gill et al., 1998; Gmitter et al., 1996; Lawson et al., 1995; Markussen et al., 1995; Quarta et al., 1999), but only in a few instances have inferred linkages with loci affecting metric or quantitative traits (QTL) been reported (Lahogue et al., 1998; Striem et al., 1996). Attempts to use MAS in fruit breeding programs are just beginning, using a few, simply inherited traits (Gardiner et al., 1999; Gill et al., 1998; Urbanietz et al., 1999).

Our objective in this paper is to compare the genetic efficiency and cost efficiency of MAS for simply inherited or polygenic traits vs. those of conventional fruit breeding schemes such as phenotypic selection among individuals and the less commonly used index selection on relatives for quantitative traits. We intend that this discussion provide fruit breeders with a framework in which to consider whether MAS will make genetic or fiscal sense in their cultivar development programs.

SPECIAL FEATURES OF FRUIT BREEDING PROGRAMS

From a plant breeder's viewpoint, fruit crops differ from most agronomic or forest crops because of a peculiar combination of features that also affect the prospect for successful use of MAS, viz., heterozygosity, asexual propagation, perennial nature, and perishability of product. Most fruit crops (with important exceptions in the genus *Prunus*, for example) have high heterozygosity in individuals and allelic richness in primary germplasm pools, which are maintained by various genetic or physical mechanisms that promote outcrossing. Propagation by asexual techniques, for testing and for cultivar release, is the norm and enables fruit breeders to conserve for horticultural exploitation all the genetic effects, additive and nonadditive, expressed in the phenotypes of superior individuals. The crops are perennials, with many featuring large plant size, long productive

period, an extended juvenile phase for seedlings, and a marketable product that cannot be assessed until a seedling is physiologically mature. Their perishable product consists largely of water and, thus, is subject to numerous interactions of genetic effects with consumer preferences, as well as with biotic and abiotic factors during preharvest and postharvest periods.

The fruit breeder requires that selected individuals exceed a certain level of performance, a culling level, for each of a host of traits desired by producers, processors, and consumers. The culling levels for most traits are usually independent of one another. Each trait is also weighted relative to its importance for the commercial success of a cultivar. Most breeders also take advantage of multiple-stage selection by emphasizing a limited number of traits when first evaluating nonreplicated seedlings and then considering the full suite of traits in advanced testing of clonally replicated genotypes.

Inheritance of traits. A few traits may have demonstrably simple inheritance because the effects of one or two genes are large compared with environmental effects. Individuals in segregating populations can usually be assigned to discrete groups corresponding directly to their genotypes. For most traits of interest, however, individuals in a segregating population cannot be assigned to discrete classes, but are distributed continuously along a quantitative scale. Termed quantitative traits, their continuous distribution is presumably due to from several (oligogenic) to many (polygenic) gene loci with effects of varying magnitude that may be small relative to environmental effects. The simultaneous selection for multiple oligogenic or polygenic traits insures that only a small proportion of individuals will have favorable alleles at a large enough number of loci to be judged superior, and necessitates evaluation of large populations to increase the probability of obtaining and identifying them.

MAS FOR SIMPLY INHERITED TRAITS

In fruit crops, the best examples of simply inherited traits are resistances to certain diseases or pests, fruit color, and plant growth habit. In some crops, such as blueberry (*Vaccinium* sp.) or strawberry (*Fragaria xananassa* Duch.), simple inheritance has been established for only a few useful traits. In other crops, such as peach [*Prunus persica* (L.) Batsch] and apple (*Malus xdomestica* Borkh.), several important traits are controlled by a single locus. These traits are often the easiest for breeders to select for, as they can be assessed readily on a single-plant basis during initial evaluation of seedlings.

The most likely candidate traits for cost-efficient MAS in fruit breeding programs (Mehlenbacher, 1995) are: 1) fruit traits that can be evaluated only after a long juvenile phase; 2) certain types of disease or pest resistance that require difficult and expensive evaluation protocols, or may be impossible to perform because the pest or pathogen is absent; 3) traits for which progeny testing is the only conventional means to confirm the presence of an allele (e.g., the pyramiding of resistance genes); and 4) simply inherited traits in species such as peach or sour cherry (*Prunus cerasus* L.) in which true backcross introgression is not hampered by inbreeding depression or self-incompatibility.

The breeder contemplating MAS for a simply inherited trait should consider several genetic and economic issues of costs, logistics, marker-gene associations, breeding practices, population sizes, and constraints of two-stage selection.

Costs. A major issue is the costs of MAS, including the development costs for markers and methods, vs. those of conventional selec-

Received for publication 30 June 2000. Accepted for publication 23 Oct. 2000. J.J.L. thanks Vincent Bus, Peter Alspach, Allan White, and Sue Gardiner of the Horticulture and Food Research Institute of New Zealand for helpful discussions and comments during the preparation of the manuscript. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper must be hereby marked *advertisement* solely to indicate this fact.

¹Corresponding author. Email address: lubyx001@umn.edu

tion techniques. For many simply inherited traits, inexpensive tests are available and markers will offer no savings.

Logistics. A mundane but critical issue is how MAS will be implemented. Are large populations to be screened? If so, is this technically feasible? To be timely, and to maximize the savings when seedlings are not carried through the juvenile phase, many breeding programs would need to evaluate as many as 10,000 to 50,000 genotypes in the 6 months before seedlings are transplanted to nurseries or orchards. If the objective is to test a much smaller set of potential parents, will any savings in cost or time that accrue to the program be substantial enough to justify implementation of the technology? Finally, if parent selection is the objective, marker selection must be at least as accurate as prediction based on phenotype, as any mistake at this stage will result in large carrying costs being incurred for undesirable plants.

Marker–gene associations. Unless the marker contains the gene itself, the success of MAS depends on marker–gene linkage disequilibrium in a population. For marker–gene associations to be useful, several criteria need to be considered. How many parents are used in a breeding program in a typical year and how are the parents mated? How robust are the marker–trait linkage relationships in the germplasm pools used by fruit breeders? Most studies of marker–gene disequilibrium in germplasm pools are based on self-pollinated species with extensive (sometimes thousands of years and generations) of domestication, breeding and selection. Once established in these crops, disequilibrium would be expected to break down very slowly because of mating scheme dynamics and intense selection.

Breeding practices. Breeding practices in self-pollinated species, or in inbred line development in outcrossing species, also facilitate the efficiency of MAS for simply inherited traits. Establishing and following linkages during inbreeding (e.g., recombinant inbred or backcross lines) are rather easy. Hence, the most practical and potentially powerful use of MAS is to facilitate introgression of simply inherited traits using backcrossing or recombinant inbreeding (Hospital et al., 1992; Tanksley and Nelson, 1996) by selecting for markers linked to the desired donor allele and against other markers in the donor parent genome.

Conversely, the mating schemes used by many breeders to foster optimum genetic advance and avoid inbreeding depression in fruit crops would tend to oppose the maintenance of marker–gene linkage relationships. Thus, an expense that is often ignored when MAS is considered in the context of a breeding program results from the need to establish and evaluate marker associations for each cross or each recombination cycle. The reasoning for this is simple. In the diverse germplasm pool of an outcrossing fruit crop there is no reason to expect a given marker genotype to be associated with a given useful gene in the germplasm at large.

Population size. Except in introgressive breeding, a breeder performing MAS for one or a few simply inherited traits will need to retain large populations with the desired marker genotype to maintain a reasonable probability of identifying one or more individuals of cultivar potential for all the other critical, quantitatively inherited traits. The number of superior individuals, with the potential to be cultivars, will be quite low in most populations because of the large number of traits for which an individual must have an outstanding genotype. The probabilities of identifying superior individuals depend on the breeder's ability to identify genetic superiority that is partially disguised by environmental effects (i.e., the concept of heritability), how genetic variation is distributed among and within families, and, not least, the number of crosses and family size (Knapp, 1998).

Constraints of two-stage selection. Two-stage independent culling (tandem selection) constrains breeding opportunities more than does single-stage selection. For example, if a breeder uses MAS in the juvenile stage to cull for one or more simply inherited disease resistance genes, the decision is made that these are absolutely necessary in every cultivar and every parent in the population. If the same initial population size is maintained prior to MAS in the juvenile stage (hoping to take a smaller set to the field and save money), then it is stochastically certain that the ultimate selection response for traits in the second stage of culling will be restricted because of genetic drift and linkage (Falconer and MacKay, 1996). Conversely, if the initial

population is increased to insure adequate population size after MAS, the cost savings will diminish because more seedlings must be produced and carried through juvenile stage MAS. In a broader perspective, this paradox reflects the conflict between obtaining cultivars rapidly and improving the breeding population. Most fruit breeders use a single population for both purposes. Absolute two-stage culling is really most appropriate for a population dedicated to cultivar improvement.

MAS FOR POLYGENIC TRAITS

Theoretical considerations. Many economically important traits in fruit crops are considered to be polygenic and quantitative in their inheritance. The theoretical basis for the inheritance of quantitative traits and their response to selection has been developed over the past 80 years (Falconer and MacKay, 1996). The inheritance may be controlled by a few to many loci with genetic effects of varying size, but these effects are often small relative to environmental and genotype \times environment interaction effects. The term heritability refers to the proportion of the phenotypic variance for a trait (Y) that can be ascribed to genetic components (broad sense heritability, H^2_Y) or to the portion of genetic variance due to additive allelic effects (narrow sense heritability, h^2_Y). Response to selection for a quantitative trait (R_Y) is a function of the intensity of selection (i), the heritability of the trait, and the amount of genetic variance for a trait ($\sigma^2_{g(Y)}$) expressed as:

$$R_Y = i_Y h_Y \sigma_{g(Y)}$$

A trait of interest (Y) may also be selected for by selecting for another trait (X), provided that genetic factors controlling the traits are located close to one another on the chromosomes (linkage) or are the same (pleiotropy). The use of MAS as proposed for quantitative traits is essentially a form of indirect selection. Selection for the trait of interest is performed by indirectly selecting on the marker phenotype, or on an index that incorporates the marker phenotype, rather than on the phenotype for the trait itself.

The expected correlated response (CR_Y) of a trait of interest (Y) to indirect selection depends on the intensity of selection for X (i_X), the heritability of X (h^2_X), the genetic variance for the trait of interest ($\sigma^2_{g(Y)}$), and the strength of the genetic association, the genetic correlation ($r_{g(YX)}$), i.e., the existence of linkage and pleiotropy, and can be expressed as follows:

$$CR_Y = i_X h_X r_{g(YX)} \sigma_{g(Y)}$$

In comparing these equations, indirect selection on X is expected to improve Y more rapidly than direct selection on Y (i.e., $CR_Y > R_Y$), assuming the same selection intensity, when

$$h_X r_{g(YX)} > h_Y$$

Essentially, this relationship suggests that if the heritability of the trait of interest (h_Y) is high, indirect selection using markers or any other trait is not likely to be more efficient because, in reality, $r_{g(YX)}$ is usually much < 1 . Considering MAS, the product of $h_X r_{g(YX)}$ will never be greater than h_Y unless the proportion of variation explained by the markers exceeds h^2_Y . This is a rare situation because as the value of h^2_Y becomes less, the power to detect useful markers diminishes drastically (Lande and Thompson, 1990; Strauss et al., 1992). Thus, selection for polygenic traits using markers alone should only be considered when it confers savings in either program cost or time to release of a cultivar.

The expected genetic efficiency of MAS depends mainly on the heritability of the trait and the proportion of the additive genetic variance explained by the marker loci (p), which, in turn, depends on a genetic correlation existing between the marker genotype and the genetic value for the trait (Falconer and MacKay, 1996; Lande and Thompson, 1990). The genetic correlation depends on the actual linkage between the markers and the QTLs and the precision with which those linkages are estimated. The precision of estimation of marker–QTL linkages depends on the number of individuals in a family and the precision with which their phenotypes for the trait are estimated, i.e., the degree to which the genetic effects are not confounded by interactions of genetic and environmental effects.

Lande and Thompson (1990) proposed a modification of indirect selection that should be more efficient than selecting on markers alone. The effectiveness depends on the alternative genotypes for a marker