

Plant Breeders' Perspectives on Improving Yield and Quality Traits in Horticultural Food Crops

James J. Luby¹

Department of Horticultural Science, University of Minnesota, 1970 Folwell Avenue, St. Paul, MN 55108

Douglas V. Shaw

Plant Sciences Department, University of California, Davis, CA 95616

Additional index words. multiple trait selection, phenotypic correlation, genetic correlation

Abstract. Breeders of horticultural food crops are usually concerned with multiple traits related to yield and quality as well as other traits such as biotic and abiotic stresses. Yield in these crops is not solely tonnage of biomass produced in the field. Rather, it is the proportion of the crop that can be harvested and brought to market in a condition and at a price acceptable to the consumer. Quality may include flavor, color, shape, size, degree of damage, nutrient levels, and traits that permit greater perceived food safety or environmental sustainability. Some traits may exhibit phenotypic associations. Traits with unfavorable associations will be of concern to the breeder if the cause is unfavorably correlated genetic effects, especially those resulting from pleiotropy. Several multiple trait selection schemes have been developed, including independent culling levels, tandem selection, and index selection. These schemes can result in improvement even for traits with unfavorable associations. However, the breeder must have a strong rationale for each trait addressed in a breeding program because each additional trait necessitates larger breeding populations and more resources. Thus, the breeder's first challenge for each crop is to determine which traits are most important and which issues are most amenable to a breeding solution.

A successful cultivar in any horticultural food crop must meet minimal criteria for numerous traits that are currently or potentially valued in the marketplace. Consumer purchases are stimulated based on "quality" traits innate to the food such as flavor, color, shape, size, degree of damage, and nutrient levels. Janick (2005) stressed that several important quality improvements such as supersweet maize types, Sugarsnap-type peas, seedless watermelon, and "Gold" pineapple caused enhanced demand or new market sectors and are some of the most important genetic breakthroughs in horticultural crop breeding. Consumers may also be attracted to products with greater perceived food safety or environmental sustainability. These characteristics may be conveyed through traceability certification programs or an ecobrand indicating where or how the food was produced. Economic sustainability, furthermore, dictates that suppliers using the cultivar must be able to produce, store, transport, market, and sell a product that appeals to the consumer and produces a profit.

Yield is seldom only about obtaining gross tonnage in horticultural food crops. Most horticultural food crops are tasty, brightly colored, biologically active packages of water. The same nutrients they provide for critical nutrition to humans also make them a marvelous medium for a host of other organisms. Thus, yield is partly about producing tonnage, but also about the proportion of the crop that can be harvested and brought to market in a condition and at a price acceptable to the consumer. Yield traits that a plant breeder must select for include biomass productivity and traditional yield components such as size and number of

harvested organs. However, genetically controlled pest, disease, and decay resistances during preharvest and postharvest stages are usually equally important. Superiority for multiple "quality" traits and "yield" traits is essential for economic sustainability in a successful cultivar.

The first challenge for a breeder is to determine which traits are most important. In a subsistence situation, focusing on content of key nutrients that are deficient in a diet that may lack quantity and diversity may be of utmost importance. The challenge in Western countries may be different. Greater nutritional benefit may be gained by substituting a few more servings of appealing, tasty, available, and inexpensive fruits and vegetables of even average nutrient content for foods with more dubious nutrient content. In this case, the breeders' efforts are better directed to traits that would make fruits and vegetables a more desirable and affordable alternative to other less nutritious products.

A related but equally important consideration for the breeder is to focus on those traits most amenable to breeding solutions. Most problems and opportunities faced by the plant breeder can be addressed genetically or horticulturally; application of horticultural solutions, where possible, can free valuable resources for improvement of traits most in need of genetic solutions. The rationale for adding an additional trait to a selection scheme must be strong. Each additional trait will mean that additional lines must be evaluated and resources expended to select for the newly added trait as well as to continue making the same rate of gain for previous traits.

The next issue a breeder encounters is how to assess quality. Ideally, the way it is assessed in the breeding program should reflect the quality that the consumer will encounter. Factors that breeders need to consider include maturity stage, location and season of production, sampling methods, and postharvest handling, storage, and pro-

cessing parameters. The example of quality components in blackberry fruit examined by Siriwoharn et al. (2004) indicates that this can be complex. They examined the variance contributed to the measurement of several quality parameters (soluble solids, phenolics, total anthocyanin, cyan 3-dioxalylglucoside, and ellagitannins) by experimental design factors, including variation among plots, subsamples within plots, multiple preparations from each subsample, and instrumental variation for multiple measurements of the same preparation. The contribution of each of these factors varied greatly depending on the parameter measured. Variation among plots accounted for almost 40% of the total variance for cyan 3-dioxalylglucoside but was negligible for the other traits. Instrumental variation among measurements accounted for almost 50% of the total variance for total anthocyanin content but was negligible for the other traits. The variation among subsamples was important for all traits, but the contribution to total variance ranged from 25% to nearly 80%. This example illustrates how breeders need to establish the environmental and experimental components that affect quality components before they can effectively hope to alter the genetic component.

ASSOCIATIONS BETWEEN TRAITS: PHENOTYPIC, GENOTYPIC, AND GENETIC CORRELATIONS

Because a successful horticultural variety requires several characteristics valued by consumers and providing economic sustainability for suppliers, every plant breeder faces the challenge of simultaneously selecting for multiple traits. Any pair of traits under selection may be independent and favorably or unfavorably associated with respect to the desired direction of selection. An association is usually detected as a product moment correlation (r_{XY}) between two traits (X and Y) in a given breeding population calculated

Received for publication 24 Mar. 2008. Accepted for publication 12 Aug. 2008.

¹To whom reprint requests should be addressed; e-mail lubyx001@umn.edu

from estimations of the covariance between traits X and Y (σ_{XY}) and the variances of X and Y (σ^2_X and σ^2_Y , respectively). That is:

$$r_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma^2_X * \sigma^2_Y}} = \frac{\sigma_{XY}}{\sigma_X * \sigma_Y}$$

When X and Y are the phenotypes of two traits within an individual (i), each is conditioned by genetic and environmental effects:

$$X_i = G_{Xi} + E_{Xi} \text{ and } Y_i = G_{Yi} + E_{Yi}$$

So, just as the phenotype of an individual has both a genetic and environmental component, a phenotypic correlation between two traits in a population can be the result of genetic (g) or environmental (e) effects or both (Falconer and Mackay, 1996), as follows:

$$\sigma_{XY} = \sigma_{g(X)g(Y)} + \sigma_{e(X)e(Y)}$$

(when G and E are independent)

$$\sigma^2_X = \sigma^2_{g(X)} + \sigma^2_{e(X)}$$

$$\sigma^2_Y = \sigma^2_{g(Y)} + \sigma^2_{e(Y)}$$

$$r_{XY} = \frac{[\sigma_{g(X)g(Y)} + \sigma_{e(X)e(Y)}]}{\sqrt{[(\sigma^2_{g(X)} + \sigma^2_{e(X)}) * (\sigma^2_{g(Y)} + \sigma^2_{e(Y)})]}}$$

Environmental correlations are the correlation of environmental effects for the two traits:

$$r_{e(XY)} = \frac{\sigma_{e(X)e(Y)}}{\sqrt{\sigma^2_{e(X)} * \sigma^2_{e(Y)}}$$

Environmental correlations result when an environmental condition (or effect) affects two distinct trait values, although their corresponding genetic effects are independent.

Genetic correlations are the correlation of “genetic” effects that contribute to phenotypes for traits X and Y:

$$r_{g(XY)} = \frac{\sigma_{g(X)g(Y)}}{\sqrt{[\sigma^2_{g(X)} * \sigma^2_{g(Y)}]}}$$

Genetic correlations, like genetic variances, can include both additive and dominance effects. Correlations that include both additive and dominance effects are usually called genotypic correlations, r_{gy} . The genotypic correlation between X and Y is relevant when considering the relationships between traits in a set of hybrids or asexually propagated genotypes. The genotypic correlation can be obtained by analysis of covariance when data for traits X and Y are collected from the set of hybrid or clonal genotypes grown in multiple environments.

Correlation of additive effects is usually called the genetic correlation or additive genetic correlation, r_a in Falconer and Mackay’s (1996) notation and r_g in others (Bernardo, 2002). The genetic correlation between X and Y can be estimated in a breeding population using controlled mating designs (e.g., diallel, nested, or factorial designs) and progeny testing of the resulting genotypes or lines in multiple environments.

The genetic correlation is the linear association of breeding values of individuals for two traits (Bernardo, 2002) and is important because it indicates the extent to which direct selection for one trait will result in an indirect response to selection for the second trait. Traits exhibiting favorable phenotypic correlations resulting from correlated additive genetic effects present the breeder with a bonus because selection for one trait will automatically result in some level of improvement for the second trait. An example of this is red color in watermelon. Breeders have selected for red flesh color to make the product more attractive to consumers, but another benefit has been enhanced content of lycopene (Perkins-Veazie et al., 2001), one of the major pigments and an important phytonutrient.

For traits with unfavorable phenotypic correlations, it is important to discern whether correlations are attributable to genetic or environmental effects. In blueberry germplasm sampled by Connor et al. (2002) over two seasons at three locations, the proportion of variation resulting from genotypes was similar to that resulting from locations. Berry weight was negatively correlated with anthocyanin and phenolic content and accounted for $\approx 31\%$ of the variation among genotypes; several outliers were identified with both large berry size and high anthocyanin and phenolic content. Connor et al. (2005) found in raspberry that the phenotypic correlations of fruit weight with antioxidant activity and total phenolic content were negative ($r = -0.33$ and -0.34 , respectively, both $P < 0.01$), but the genetic correlations were not significant ($r = 0.07$ and 0.01 , respectively, both $P > 0.05$). These examples illustrate that just because an unfavorable phenotypic correlation is observed between a quality trait and a yield-related trait, it is a mistake to simply assume that it is the result of genetic factors.

The manner in which phenotypic correlations are determined can have an effect. Unpublished data from the strawberry breeding program of DVS indicates that sugar content of the fruit is negatively correlated ($r = -0.39$, $P < 0.01$) with yield when examined during individual 3-week periods of the harvest season, but was uncorrelated over the whole harvest season ($r = 0.02$, $P > 0.05$). Farnham et al. (2004) reported that the concentration glucoraphanin, a chemopreventive compound, was not significantly correlated ($r = 0.02$, $P > 0.05$) with broccoli head weight, a yield component, whereas total glucoraphanin per head was positively correlated ($r = 0.50$, $P < 0.01$).

Genetic correlations can be caused by two factors (Falconer and Mckay, 1996). The first is pleiotropy in which a gene(s) has an effect on more than one trait. Pleiotropy might be the result of biochemical products that affect more than one aspect of a plant phenotype, a physiological limitation such as earliness of harvest and yield. A second cause of genetic correlations is gene linkage or, more accurately, linkage disequilibrium: the nonrandom

association between alleles at different loci. Linkage disequilibrium can be generated whenever two parents are mated in a breeding program and will result in genetic correlations when genes conditioning two physiologically distinct traits are associated. Correlations resulting from disequilibrium are expected to be transient unless maintained by selection, and this may explain why genetic correlations are notorious for lack of consistency over generations and among tests. They can be reduced with repeated cycles of meiosis and in plant breeding populations, this can be accomplished with several generations of random intermating.

A breeder would have little reason to expect strong unfavorable genetic correlations between yield and most quality traits a priori provided that selection is conducted in a population that is maintained by frequent cycles of random mating. Situations that involve germplasm infusion or hybridization among populations frequently generate negative genetic correlations between pairs of traits that differ in value between the parent sources and demand special caution. Regardless of breeding method, the starting point for any breeder would be to understand these correlations. Correlations between traits will especially concern a breeder when they are unfavorable, strong, and genetic as a result of pleiotropy. However, just because two traits exhibit an unfavorable phenotypic correlation, an unfavorable genetic correlation cannot be assumed.

Randomization and replication of genotypes in a single testing environment can remove the correlation between genetic and environmental effects but will only reflect the microenvironmental correlation between two traits. The phenotypic and genotypic correlations between two traits can be estimated directly and easily from analysis of covariance of an experiment with genotypes grown in multiple environments. The genetic correlation between traits can only be estimated from progeny tests in multiple environments. Genetic correlations are prone to larger SES compared with phenotypic correlations (Reeve, 1955; Robertson, 1959) because they are estimated from multiple variance and covariance components, each having an associated error. Thus, large numbers of progeny in multiple environments are needed to obtain precise estimates, and the expense of obtaining sufficiently large data sets may explain why they are seldom reported by horticultural researchers and breeders. In multiple environment tests with plants, the contribution of genotype \times environment interactions to covariance may also be important and these contributions can also be estimated (Aastveit and Aastveit, 1993).

BREEDING METHODS TO ADDRESS MULTIPLE TRAITS AND UNFAVORABLE GENETIC CORRELATIONS

Improvement programs always must consider more than one trait. The traits may not be given equal emphasis, however. Certain

traits may be more economically important than others, traits for which consumers are willing to pay a premium. Alternatively, some traits may receive more attention because they are easier to improve as a result of higher heritability or rapid and inexpensive testing procedures. Plant breeders have developed several methods to deal with the task of multiple trait selection.

Tandem selection attempts to improve a breeding population for several traits by focusing on selection for one trait at a time in alternating generations or cycles of selection. Developing separate lines for each trait is a special case of tandem selection. Tandem selection is of little use in fruit crops and other perennial species with long generation intervals. In these crops, complementary mating is often used and has a similar objective to tandem selection. In complementary mating, a parent selected because it was exceptional for one trait but deficient for a second is mated with a parent that is exceptional for the second trait although it may be deficient for the first trait with the objective of identifying offspring that are exceptional in both traits.

Independent culling involves selection for multiple traits in a specified order within a single generation and from a single initial population. Bernardo (2002) suggests that many plant breeders often use intuitive, if not explicit, independent culling levels when selecting a parent for population improvement or in inbred and hybrid development programs. The breeder must establish minimum standards (culling levels) for each trait. Each tested genotype must meet the standard for each trait to be selected for the next breeding cycle. A genotype can be culled based on a single trait without waiting for data on other traits and, therefore, selection can be conducted in stages to facilitate program efficiency. The weighting of importance for different traits is achieved by allocation of selection intensity. Traits considered more important are selected with greater intensity, e.g., an individual must be in the top 10% of the population to be retained for the most important trait but in the top 25% of the population for a trait of less importance. When genetic correlations are unfavorable, selection intensity can be relaxed. A less strict culling level for one trait will help ensure that there is sufficient variation for an unfavorably correlated trait.

Selection using independent culling levels dictates that the breeder maintains a sufficiently large population size after each successive stage of culling to ensure that sufficient variation remains for traits in subsequent culling. In cultivar development, preliminary screening followed by more rigorous levels of testing is a very common practice. Large numbers of genotypes are first screened in low-precision test followed by intensive testing of a smaller number of promising items. This process differs from simple independent culling because the initial selection stage reduces genetic variability. When genetic correlations are unimportant, the order of

culling is not important to gain but may be very important to the economics of the breeding program. The ability to divide culling into two economically or operationally efficient stages is probably the most important advantage of independent culling. For example, culling for disease resistance might be done at the juvenile stage in a greenhouse screening of seedlings, and only genotypes not culled at the juvenile stage are further evaluated for yield and quality traits in later field tests.

Index selection involves simultaneous selection of multiple traits in the same generation. The breeder creates a single new "trait," the index, which is a function of the multiple traits that are under selection. Selection among genotypes or lines is then performed based solely on the index value rather than the values of the individual traits.

Index selection is essentially a method for weighting the individual traits based on their perceived importance and the opportunity for improvement. Index selection is based on the concept that each genotype has an aggregate genetic value that is a function of its performance for the multiple traits being considered. This aggregate genetic value, H , can be defined as:

$$H = v_1G_1 + v_2G_2 + \dots$$

where v_i = economic value and G_i = genetic value for trait i . In a given cycle of selection, the economic value is constant, but genetic value differs among individuals. Over time, the breeder may alter the economic value based on changes in the importance of traits in the marketplace. Mean genetic value for a trait in a breeding population is expected to improve over successive cycles of selection, although the rate of improvement will depend on not only on the economic value assigned to the trait, but also on the genetic variation and heritability of the trait in the population.

Many horticultural breeders use an intuitive index when making selections, even if they do not develop an explicit index. Index selection is theoretically the most efficient method for improving crop value given the concept of aggregate genetic value (Bernardo, 2002). A major feature of index selection in dealing with unfavorably correlated traits is that an individual that is outstanding for one trait but average for another will be retained; this is not the case for independent culling.

In summary, a breeder of a horticultural food crop can always expect to deal with the necessity of improving multiple traits. These will likely include traits related to yield and quality as well as tolerance of diseases, pests, and abiotic stresses. Many of these traits will exhibit no association. For the breeder, those that exhibit an unfavorable association resulting from correlated genetic effects will be most problematic. Several breeding methods have been described that can potentially manage the effects that unfavorable genetic correlations might have on correlated responses to selection with multiple traits. The independent culling levels method is often practical and is predicted to be more

effective than tandem selection for improving multiple traits, whereas index selection is usually the most efficient (Hazel and Lush, 1942). Bernardo (2002) concluded from reviewing studies in many agronomic crops that any judiciously applied selection index is useful for simultaneous improvement of multiple traits. Regardless of the method used, sufficiently large breeding populations need to be used. A sufficiently large population size allows the detection of rare genotypes resulting from recombination when genetic correlation is the result of loci in linkage disequilibrium. If these loci have major effects and are well characterized, use of DNA marker-assisted selection may facilitate multiple trait selection (Luby and Shaw, 2001). Unfavorable genetic correlation resulting from pleiotropic effects is a more difficult situation that can be overcome only when alternative biochemical pathways or physiological processes controlled by other genes are brought together in an outlier genotype. Novel germplasm or parental combinations may be necessary to accomplish this along with evaluation of large population sizes to allow detection of outliers.

Literature Cited

- Aastveit, A.H. and K. Aastveit. 1993. Effects of genotype-environment interactions on genetic correlations. *Theor. Appl. Genet.* 86:1007-1013.
- Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN.
- Connor, A.M., J.J. Luby, C.B.S. Tong, C.E. Finn, and J.F. Hancock. 2002. Genotypic and environmental variation in antioxidant activity, total phenolic content, and anthocyanin content among blueberry cultivars. *J. Amer. Soc. Hort. Sci.* 127:89-97.
- Connor, A.M., M.J. Stephens, H.K. Hall, and P.A. Alspach. 2005. Variation and heritabilities of antioxidant activity and total phenolic content estimated from a red raspberry factorial experiment. *J. Amer. Soc. Hort. Sci.* 130:403-411.
- Falconer, D.S. and T.F.C. Mackay. 1996. Introduction to quantitative genetics, Ed 4. Longmans Green, Harlow, Essex, UK.
- Farnham, M.W., P.E. Wilson, K.K. Stephenson, and J.W. Fahey. 2004. Genetic and environmental effects on glucosinolate content and chemoprotective potency of broccoli. *Plant Breed.* 123:60-65.
- Hazel, L.N. and J.L. Lush. 1942. The efficiency of three methods of selection. *J. Hered.* 33:393-399.
- Janick, J. 2005. Horticultural plant breeding: Past accomplishments and future directions. *Acta Hort.* 694:61-65.
- Luby, J.J. and D.V. Shaw. 2001. Does marker-assisted selection make dollars and sense in a fruit breeding program? *HortScience* 36:872-879.
- Perkins-Veazie, P., J.K. Collins, S.D. Pair, and W. Roberts. 2001. Lycopene content differs among red-fleshed watermelon cultivars. *J. Sci. Food Agr.* 81:983-987.
- Reeve, E.C.R. 1955. The variance of the genetic correlation coefficient. *Biometrics* 11:357-374.
- Robertson, A. 1959. The sampling variance of the genetic correlation coefficient. *Biometrics* 15: 469-485.
- Siriwoharn, T., R.E. Wroldstad, C.E. Finn, and C.B. Pereira. 2004. Influence of cultivar, maturity, and sampling on blackberry (*Rubus* L. hybrids) anthocyanins, polyphenolics, and antioxidant properties. *J. Agr. Food Chem.* 52:8021-8030.