Genetic and Environmental Variation for Tomato Flesh Color in a Population of Modern Breeding Lines

Erik J. Sacks and David M. Francis
Department of Horticulture and Crop Science, The Ohio State University, Ohio Agricultural Research and Development Center, 1680 Madison Ave., Wooster, OH 44691-4096.

Additional index words. CIELAB, Lycopersicon esculentum

ABSTRACT. The genetic and environmental variation for flesh color of tomato (Lycopersicon esculentum Mill.) fruit was quantified using 41 red-fruited breeding lines, open-pollinated cultivars, and hybrids that are representative of the diversity of tomatoes grown for whole-leaf processing in the midwestern and eastern United States and Ontario, Canada. Objective color measurements were made for 2 years from replicated experiments with 2 to 4 blocks per year. Genotypes differed significantly in lightness value (L*), saturation (chroma), and hue angle. Variation within fruit and among fruit in plots accounted for more than 75% of the environmental variation for the color traits. The crimson locus (og c ) accounted for less than one-third of the variation in fruit color among genotypic means, and explained 18% to 27% of the genotypic variation for L*, chroma, and hue. Estimates of variance components were used to develop sampling strategies for improving selection efficiency. Genotypes were identified that may be useful for studying genetic differences that lead to quantitative variation for fruit color in red-fruited populations of tomato.

Color is a defining attribute of quality for fresh and processed tomatoes. Given the predominance of red fruit in tomato commerce, our breeding goal is to obtain genotypes that consistently produce fruit with color that is saturated, dark, and red. Breeding red tomatoes of improved color from modern breeding lines requires that we be able to distinguish between genetic and environmental effects in populations exhibiting continuous phenotypic variation. Moreover, the identification of breeding lines that differ in color is a required first step towards understanding how genetic differences lead to these small, but horticulturally important, differences in phenotype.

Pigmentation in tomato is controlled by many genes (Stommel and Haynes, 1994; Tigchelaar, 1986). Nine classically defined genetic loci (with a total of 15 alleles) have large effects on the flesh color of ripe tomatoes. These loci include: old gold crimson (og c ) and its allele Beta-carotene; apricot; Delta; diospyros; green flesh; Green ripe; high pigment-1; high pigment-2; dark green; Intense pigment; modifier Beta-carotene; red color in yellow fruit; sherry; tangerine; and yellow flesh (Jenkins and Mackinney, 1953, 1955; Kerr, 1958; Rick and Chetelat, 1993; Thompson et al., 1967; Tomes et al., 1953, 1956; Van Tuinen et al., 1997; Young, 1956). Many of these genes affect fruit color by controlling the quantities and/or types of pigments produced (Khudairi, 1972). Unfortunately, many of these classically defined genes are of no direct use for breeding red tomatoes. Genotypes with the high pigment genes (including dark green) increase total carotenoid pigments and improve fruit color, but typically exhibit undesirable characteristics such as poor germinability and brittle stems (Jarret et al., 1984; Thompson, 1961). Thus undesirable pleiotropic effects have slowed the use of high pigment genes. In contrast, og c has been used extensively in processing tomato breeding programs for the midwestern and eastern United States. This gene works through a biochemical mechanism that is distinct from the high pigment genes, as it increases lycopene content and reduces β-carotene content (Thompson et al., 1965).

Recently alleles from wild species have been identified that increase red color when transferred into cultivated tomato (Bernacchi et al., 1998; Tanksley et al., 1996). These alleles have been identified using a combination of quantitative trait locus (QTL) analysis, DNA-based genetic markers, and an advanced backcross breeding strategy. Five loci with a positive effect on fruit color were identified in L. hirsutum Dunal accession LA 1777 (Bernacchi et al., 1998) and three in L. pimpinellifolium (B. Juss.) Miller accession LA 1589 (Tanksley et al., 1996). Several of these loci map to positions in the genome lacking classically defined loci that affect color. Thus, in addition to the classically defined major genes, other genes that affect fruit pigmentation and color probably exist and can be defined using quantitative methods.

Evidence exists for additional genes that have a quantitative effect on fruit color and pigment content within L. esculentum. For example, the amounts of predominant carotenoids vary widely in the ripe fruit of three old cultivars, Campbell 146, Eastern States 24, and Rutgers. For these three varieties, pigment content ranged from 19% to 42% colorless phytoene and phytofluene, 1% to 2% pink ζ-carotene, 4% to 6% orange β-carotene, and 50% to 76% red lycopene (Thompson et al., 1967; Tomes, 1963). Raymundo et al. (1976) found that ripe, red fruit of ‘Summer Sunrise’ contained 1.7% greenish-yellow ζ-carotene and 0.1% yellow neurosporene in addition to 17% phytoene and phytofluene, 0.5% γ-carotene, 6.9% β-carotene, and 73.7% lycopene. Thompson et al. (1967) noted that the genetic background of ‘Campbell 146’ and ‘Eastern States 24’ influenced the expression of og c in F1 and backcross populations derived from crosses with ‘High Crimson’ (og/og). Thus it is likely that genes or alleles other than og c may be useful for improving color.

Given that environment can obscure genetic differences (McCollum, 1956), breeding progress will require estimates of fruit color that are precise enough to distinguish genotypic
differences among breeding lines of red-fruited tomatoes. One obstacle to improving tomato color through breeding is a lack of knowledge on how best to sample. In this study, we quantified sources of genotype × environment variation and devised sampling strategies that should improve detection of small differences among genotypes for flesh color. Our goal was to develop a strategy that allows for increased efficiency in breeding tomatoes suited for a whole-pee product. We identified modern, red-fruited tomato lines that differ in color under horticulturally relevant environments. In addition, genotypes lacking og* were identified that produced dark, red, highly saturated fruit. Thus the genetic control of color in some of the lines is known to be distinct and, therefore, may provide useful material for studying the genetic mechanisms underlying quantitative variation for color within adapted germplasm.

Materials and Methods

**Plant materials.** A sample of 41 modern breeding lines and cultivars was chosen to represent a cross-section of the germplasm of processing tomatoes typically grown in the Midwestern and Eastern United States and Ontario, Canada (Table 1). Seed was planted in flats on 6 Apr. 1995 and 3 Apr. 1996. Seedlings were grown inside a greenhouse and transplanted to the field 6 to 8 weeks after planting. Transplants were spaced 0.3 m apart on raised beds with 1.54 m between the beds. Field trials were conducted at The Ohio State University’s Vegetable Crops Branch in Fremont, Ohio, which is located in an area of commercial tomato production. Production practices were as recommended for commercial growers (Prechuer, 1997).

A Johnson mechanical harvester was used for once over harvest and fruit were sorted by hand into ripe, green, and cull categories. The ripe fruit were graded at 98% to 100% usable by U.S. Department of Agriculture inspectors and provided the raw material for color analysis.

**Color measurement.** Five to thirteen ripe fruit were measured from each plot. Two objective measurements in CIELAB color space (Commission Internationale de l’Eclairage, 1978) were obtained for each fruit with a CR-100 colorimeter (Minolta, Ramsey, N.J.). The CR-100 colorimeter measures the red, green, blue, and total amount of light reflected from an object using an 8-mm-diameter measuring area, a d/0° illuminating and viewing geometry, and illuminant C. CIELAB has been used for objective color measurements of horticultural commodities such as gerbera (Gerbera jamesonii H. Bolus ex Hooker) flowers (Tourjee et al., 1995) and strawberry (Fragaria xanassa Duch.) fruit (Sacks and Shaw, 1994). In the CIELAB color space, +a* is the red direction, −a* is the green direction, +b* is the yellow direction, and −b* is the green direction. The color space is three-dimensional, where the third axis, L*, represents white to black and the a*−b* plane may be visualized as a color wheel that is lighter or darker depending on the level of L*. Lower L* values represent a darker color. Chroma (C*), a measure of saturation, was calculated with the formula: (a*² + b*²)½. A minimum color CIELAB difference of 1 unit will be perceptible to a human observer depending on the L* value, background color, and lighting; differences in the 1 to 2 unit range are noticeable to most observers (E. Carter, Minolta, personal communication; Berger-Schunn, 1994). Hue-angle, in degrees, is a measure of an object’s color in the a*–b* plane and was calculated as (180/p) * cos⁻¹(α*/C*). For the positive values of b obtained, Perception of hue angle differences will depend on the chroma with differences more detectable at higher chroma. In general, and based on the assumption that there are ≈160 distinguishable hues, a hue angle difference of 2.5 is detectable (Hardin, 1990).

Fruit flesh was exposed for measurement by cutting the proximal end of the tomato transversely with a sharp knife, such that only the pericarp at the top of the fruit was visible (Francis et al., 2000). The color of the gelatinous placental tissue was not measured, but two measurements were made on opposite sides of the exposed fleshy surface.

**Experimental design.** A randomized complete block design was used. Each block had one plot per genotype. Plots were 6.17 m in length consisting of 20 plants per plot and were harvested as a unit. There were four blocks of 41 genotypes in 1995, and two blocks with all 41 genotypes in 1996. Two additional blocks of 14 of the genotypes (for a total of 4 blocks) were planted in 1996. The experimental design included two levels of subsampling: among fruit in plots, and measures within fruit. In the statistical analysis plot is equivalent to genotype by block within year (G × BY); within plot variation was fruit within genotype by block by year (Fruit×GBY); and within fruit variation was equivalent to error.

**Statistical methods for main analysis (41 genotypes and 6 blocks).** Mixed model analyses of variance were performed with the SAS (1992) procedure MIXED. Analyses were conducted for the dependent variables: L*, chroma, and hue angle. Genotype (G) was considered a fixed effect. The random effects were year (Y), block within year (BY), G × Y, G × BY, and fruitGBY. Degrees of freedom were calculated via the Satterthwaite option (Neter et al., 1990). Standard errors for proportions of environmental variance were estimated by the method of Dickerson (1969). Means were obtained with the lsmeans statement. Linear correlations of genotype means were calculated.

We calculated the number of tomato fruit (Fr) per block and year needed to estimate with 90% or 95% confidence a genotype’s flesh color within an upper-lower interval (I). These calculations were made for L*, and chroma in CIELAB units or degrees for hue angle and were based on taking two sample measurements ‘S’ per fruit from ‘B’ blocks in each of ‘Y’ years. Estimates of Fr were derived from the formula for confidence intervals about a sample mean (Steel, et al. 1997), I = 2 × t × s/n ½ , based on the assumption that the variance of a genotype’s mean is, σ² = σ²/Y + σ²/GBY + σ²/FRBY + σ²/SFRBY. Thus for our data set, (I/2)² = f² × σ²/Y + z² × σ²/GBY + z² × σ²/FRBY + z² × σ²/SFRBY. Solving for Fr, Fr = [(z²/SBY) × (σ²/Y) + S²fr]/[(I/2)² – (z² × σ²/GBY / Y) – f² × σ²/Y]. Values of f from the t distribution were based on 40 degrees of freedom associated with σ²/y. For all other interactions the df were so large that critical values approximate those for df = infinity, hence z values were chosen over t values. For the 95% interval, z was set to 1.96 and t was 2.021. For the 90% interval, z was set to 1.645 and t was 1.68.

To determine the effect of the crimson gene, the presence of the og* allele was confirmed by both pedigree and detection of gold flower color under cool greenhouse conditions during winter 1998. An analysis of variance was conducted on the 41 genotypes with the SAS (1988) procedure GLM to study the effects of three levels of crimson (og/og, og/+), and +/+ on L*, chroma, and hue. The data were the means for each genotype calculated in the main analysis (Table 1) and subanalysis.

**Subanalysis (14 genotypes and 8 blocks).** A data set that included a subset of 14 genotypes from the main analysis (Table 1), but with two additional blocks and =20 additional fruit per genotype, was studied. Statistical procedures were analogous to those used in the main analysis.
Results and Discussion

The range of means we observed (Table 1) indicates that there were appreciable differences in color among the genotypes that would be detectable by most human observers. The objective color data in CIELAB color space reported in Table 1 is related to the tomato industry standard of a*/b* ratio by the regression equation hue angle = (a*/b*)(–35.05) + 79.70 for the values of hue angle obtained in this study (R² = 0.99, P ≤ 0.001). This equation is not valid for hue angles outside of the range reported in Table 1 because linear functions only approximate the trigonometric function in a narrow range of a*/b* ratios (Little, 1975). The large range for hue angles (measured in degrees) among the genotypes confirms that the fruit of some genotypes was more orange than that of other genotypes. The objective color data were consistent with our subjective observations that lower hue angles correspond to more red in the fruit and higher hue angles correspond to more orange (45°) or yellow (90°) fruit.

In the main analysis of variance, differences among genotypes were significant for L* (P = 0.015) and marginally nonsignificant.
chroma ($P = 0.053$), and hue angle ($P = 0.051$). In the subanalysis containing additional replication, genotypic differences were highly significant for all three color traits ($P = 0.001$ for $L$; $P = 0.004$ for chroma; and $P \leq 0.001$ for hue angle). Thus, the additional sampling of blocks and fruit in the subanalysis increased our ability to detect differences among the genotypes. Genotypic differences in hue may have resulted from the differences in types of pigments, in ratios of pigments, or both. Differences in $L^*$ and chroma probably reflected different pigment concentrations. Given that analysis of pigments is expensive, this study’s identification of genotypes that differ in fruit color should improve the efficiency of future studies of pigment composition.

A strong positive linear correlation was observed between $L^*$ and hue angle ($R^2 = 0.83$, $P \leq 0.001$). Thus, the genotypes with lighter fruit color (higher $L^*$ values) tended to be a more yellow orange than the genotypes with darker fruit color. Ranges of genotypic chromas were similar along the entire range of hues. Evaluation of data plots did not suggest nonlinear relationships for any pair of color traits.

Table 3. Number of tomato fruit per block and year needed to estimate with 90% or 95% confidence the lightness ($L^*$) and chroma of a genotype’s flesh within an interval (upper – lower) of CIELAB units.

<table>
<thead>
<tr>
<th>No. years</th>
<th>No. blocks</th>
<th>No. 3 units</th>
<th>90%</th>
<th>95%</th>
<th>90%</th>
<th>95%</th>
<th>90%</th>
<th>95%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chroma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Variance components, standard errors (SE) of variance components, proportions of environmental variance (PEV), and standard errors of proportions for five flesh color traits of tomato.

<table>
<thead>
<tr>
<th>Color trait</th>
<th>Source</th>
<th>$\sigma^2$</th>
<th>SE of $\sigma^2$</th>
<th>PEV</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>L*</td>
<td>Year (Y)</td>
<td>0.69$^{**}$</td>
<td>1.11</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Block\Y (BY)</td>
<td>0.05$^{**}$</td>
<td>0.14</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>G x Y</td>
<td>1.30$^{**}$</td>
<td>0.80</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>G x BY</td>
<td>2.50$^{***}$</td>
<td>0.66</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Fruit\GBY</td>
<td>15.71$^{***}$</td>
<td>0.90</td>
<td>0.43</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>16.46$^{***}$</td>
<td>0.55</td>
<td>0.45</td>
<td>0.02</td>
</tr>
<tr>
<td>Chroma</td>
<td>Year (Y)</td>
<td>1.02$^{**}$</td>
<td>1.52</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Block\Y (BY)</td>
<td>0.01$^{**}$</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>G x Y</td>
<td>1.33$^{**}$</td>
<td>0.49</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>G x BY</td>
<td>1.14$^{***}$</td>
<td>0.25</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Fruit\GBY</td>
<td>3.10$^{***}$</td>
<td>0.32</td>
<td>0.19</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>9.31$^{***}$</td>
<td>0.31</td>
<td>0.59</td>
<td>0.02</td>
</tr>
<tr>
<td>Hue</td>
<td>Year (Y)</td>
<td>20.19$^{***}$</td>
<td>29.19</td>
<td>0.11</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Block\Y (BY)</td>
<td>0.27$^{**}$</td>
<td>0.55</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>G x Y</td>
<td>6.74$^{*}$</td>
<td>3.35</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>G x BY</td>
<td>5.89$^{***}$</td>
<td>2.50</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Fruit\GBY</td>
<td>89.74$^{***}$</td>
<td>4.36</td>
<td>0.49</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>58.78$^{**}$</td>
<td>1.97</td>
<td>0.32</td>
<td>0.01</td>
</tr>
</tbody>
</table>

$^{**}$, $^{***}$, $^{****}$ Nonsignificant or significant at $P \leq 0.01$, 0.001, or 0.0001, respectively.
As expected (Thompson et al., 1967), the crimson allele had a large effect on fruit color (Table 1). Significant differences between og/log and +/+ were observed for L*, chroma, and hue angle. Lines that were og/log tended to have fruit that were darker, more saturated and less orange than lines that were +/+.

The mean for og/+ was intermediate to, but not significantly different from, the means for og/log and +/+.

The mean for og/+ is consistent with the observation that og* is codominant (Thompson et al., 1967). The lack of significance for differences between the og heterozygotes and the two classes of homozygotes is probably the result of small sample size; there were only three heterozygotes in the population.

The crimson locus did not account for all of the genotypic variation. Some +/+ lines (e.g., O9442) were among the darkest-fruited and some og/log lines (e.g., O9439) had considerably lighter fruit color than many +/+ lines (Table 1). Values of $R^2$ from the analysis of variance (SAS, 1988) were 0.18 for chroma, 0.26 for hue, and 0.27 for L*, suggesting that og* explains 18% to 27% of the observed variation between genotypes. When data from only the +/+ lines were analyzed, significant genotypic variation was detected for chroma ($P \leq 0.05$). Thus genes other than og* are probably contributing to genetic variation in this population of adapted breeding lines, cultivars, and hybrids.

Variation components indicated that variation within fruit and among fruit in plots accounted for more than 75% of the environmental variation for the five color traits (Table 2). Most of the within-fruit variation is due to yellow shoulder disorder (YSD) which is characterized by sectors of yellow or green tissue under the peel (Francis et al., 2000). The occurrence of this disorder resulted in an inflated error variance because of the large difference in CIELAB values between green and red tissue relative to the more subtle differences between red and orange-red fruit associated with genotypes (Table 1). In addition, the microenvironment of a ripening fruit and relative maturity may have had an effect on within-plot fruit color. Artificial shading, and temperatures exceeding 32 °C, have been associated with low lycopene content and poor color in tomatoes (McCollum, 1956). Although these sources of variation may obscure genetic differences, the variation associated with YSD and microenvironment is normal within the context of tomatoes grown in the field for once-over machine harvest. Variance components for Y, BY and G × Y were nonsignificant or small, but those for G × BY were significant and accounted for 3% to 7% of the environmental variation (Table 2).

Depending on the desired degree of accuracy and the resources available for measuring fruit, we can choose from a number of feasible sampling strategies (Table 3, Table 4). For most color traits, the genotypic extremes of our population could be distinguished from each other by sampling a modest number of fruit (four to 41 per block) from only four blocks in 1 year (Table 3, Table 4). A higher degree of precision (four-unit intervals) is obtainable for L* and chroma with 95% confidence by sampling 24 fruit in each of six blocks for at least 2 years. Our relatively low estimates of $\sigma^2_{Y}$ and $\sigma^2_{GL}$ were obtained from only 2 years of sampling. If the true values of these variance components are actually larger than those estimated, then sampling over more years would be needed to obtain highly precise estimates of a genotype’s fruit color. Any differences among locations may also contribute to color variation. The fruit samples for this study were drawn from a location with a high incidence of YSD (Francis et al., 2000), and we have probably overestimated the within-fruit and within-plot variation relative to environments that are conducive to uniform fruit color. If, through the addition of further locations, $\sigma^2_{GL}$ and related interactions were large and the within-fruit and within-plot variance components were not affected, then more sampling would be required.

We have determined that some modern breeding lines of tomato differ for color of fruit flesh under agriculturally relevant environments. The sampling strategies presented in Table 3 and Table 4 should help tomato breeders improve fruit color via selection. Future comparisons of carotenoid composition in poor and good-colored lines, with and without og*, would be informative. By comparing the relative quantities of carotenoid pathway intermediates among breeding lines that exhibit small differences in fruit color, we may be able to focus our attention on the enzymatic and genetic differences that contribute to quantitative variation.

**Literature Cited**


Francis, D.M., S.A. Barringer, and R.E. Whitmoyer. 2000 Ultrastructural characterization of yellow shoulder disorder in a uniform ripening

---

**Table 4.** Number² of tomato fruit per block and year needed to estimate with 90% or 95% confidence the hue angle of a genotype’s flesh within an interval (upper – lower) of degrees in the CIELAB color space.