J. Amer. Soc. Hort. Sci. 133(4):579–586. 2008.

# Tomato Analyzer-color Test: A New Tool for Efficient Digital Phenotyping

**Audrey Darrigues, Jack Hall, Esther van der Knaap, and David M. Francis[1]**
*Department of Horticulture and Crop Science, The Ohio State University, Ohio Agricultural Research and Development Center, 1680 Madison Avenue, Wooster, OH 44691*

**Nancy Dujmovic and Simon Gray**
*Department of Mathematics and Computer Science, College of Wooster, 1189 Beall Avenue, Wooster, OH 44691*

**Additional index words.** software, objective color, *Solanum lycopersicum*, genetic analysis

**Abstract. Measuring plant characteristics via image analysis has the potential to increase the objectivity of phenotypic evaluations, provides data amenable to quantitative analysis, and is compatible with databases that aim to combine phenotypic and genotypic data. We describe a new tool, which is implemented in the Tomato Analyzer (TA) software application, called Color Test (TACT). This tool allows for accurate quantification of color and color uniformity, and allows scanning devices to be calibrated using color standards. To test the accuracy and precision of TACT, we measured internal fruit color of tomato (*Solanum lycopersicum* L.) with a colorimeter and from scanned images. We show high correlations ($r > 0.96$) and linearity of L*, a*, and b* values obtained with TACT and the colorimeter. We estimated genotypic variances associated with color parameters and show that the proportion of total phenotypic variance attributed to genotype for color and color uniformity measured with TACT was significantly higher than estimates obtained from the colorimeter. Genotypic variance nearly doubled for all color and color uniformity traits when collecting data with TACT. This digital phenotyping technique can also be applied to the characterization of color in other fruit and vegetable crops.**

Digital phenotyping aims to accurately describe a trait based on analysis of electronic images. Computer-based analysis of objects from digital images has the potential to increase the objectivity of data collection while reducing subjective characterization that is typically prone to bias. There are a number of computer image acquisition and analysis techniques for color in foods such as apple [*Malus ×domestica* Borkh (Leemans et al., 2002; Li et al., 2002)], banana [*Musa cavendishii* L. (Mendoza and Aguilera, 2004)], chicory [*Cichorium intybus* L. (Zhang et al., 2003)], as well as seed analysis (Granitto et al., 2002; Sako et al., 2001; Shahin and Symons, 2001) and meat (O'Sullivan et al., 2003; Tan, 2004). Color image analysis is also prevalent in floricultural crops such as lisianthus [*Eustoma grandiflorum* Grise. (Yoshioka et al., 2006)], and begonia [*Begonia ×tuberhybrida* Voss. (Lootens et al., 2007)]. Digital color analysis is also performed in plant pathology to quantify lesions on diseased leaves (Kwack et al., 2005).

Objective and systematic descriptions, trait ontologies, are being developed in the plant sciences for database retrieval and archiving (reviewed in Brewer et al., 2006; Ilic et al., 2007). This trend is stimulated, in part, by a desire to link trait descriptions to the growing databases of sequence information. Tomato has become the prominent model horticultural crop for studies in genetics and genomic sciences. With extensive resources, including 357,477 expressed sequence tags (National Center for Biotechnology Information, 2008) and a genome sequencing project focused on euchromatin (Mueller et al., 2005), there is promise for research that seeks to integrate emerging sequence resources with phenotypic variance. Fulfilling this promise will require extensive data for the traits studied. Immortal populations (e.g., recombinant inbred populations and inbred backcross populations) consist of nearly homozygous lines that preserve the genetic integrity of mapping populations. These populations serve as a resource and allow for replication of experiments and extensive analyses from different laboratories. Phenotypic and molecular characterization of such populations can be stored in public databases for use by other researchers. The Tomato Analyzer (TA) software application was developed to facilitate the collecting and sharing of data related to fruit size and shape and the identification of genes that contribute to quantitative variation in morphology (Brewer et al., 2006). We describe a new module implemented in TA that can accurately collect objective data for color from digital images.

Measuring color from digital images requires standardization and interpretation because digital devices use a color space that is not standardized, is nonlinear, and may vary between hardware devices and software applications. In the Red Green Blue (RGB) color space, each pixel is represented in the computer or interface hardware as values of red, green, and blue. In contrast, color spaces such as CIELab were designed to approximate human perception of color [Commission Internationale de l'Éclairage (CIE), 1978]. CIELab color space is a

Received for publication 28 Nov. 2007. Accepted for publication 8 Apr. 2008. We thank David Sullivan (Department of Mathematics and Computer Science, College of Wooster, OH) for his help in designing and optimizing the software application for TACT. We acknowledge Bert Bishop (Statistics and Computing Services, OARDC/OSU, Wooster, OH) for his suggestions on the analysis of the IBC population. We appreciate the help from Troy Aldrich for establishing field plots.
The development of Tomato Analyzer and its Color Test was supported by the National Science Foundation (grant no. DBI 0227541 to E.V.D.K.). A.D. was supported through the Ohio Agricultural Research and Development Center Research Enhancement Competitive Grants Program and through funds from the Mid-America Food Processors Association and from USDA/NRI grant 2004-01575.
[1]Corresponding author. E-mail: francis.77@osu.edu.

reference standard and is most commonly used for measuring object color. Color data collected in the dimensions of the CIELab color space can be archived and used for quantitative analysis of color.

Color holds an important economic role in horticultural crops. For fresh or processed products, color is one of the primary determinants of quality, along with texture, size, and flavor (Picha, 2006). In the case of tomato, color and color uniformity contribute to quality. The presence of yellow shoulder disorder (YSD) is a major quality constraint. YSD is a blotchy ripening disorder that is characterized by discolored regions under the epidermis of mature fruit. Cells from YSD tissue are smaller and more randomly organized, and the conversion from chloroplasts to chromoplasts is altered (Francis et al., 2000). Variation for YSD within fruit and among fruit in plots explained more than 75% of the variation for color (Sacks and Francis, 2001). Color disorders are also an economic problem. U.S. Department of Agriculture (USDA) processor grades are largely determined by the amount of off-color tissue in products (USDA, 2005). To improve the output of high-quality product, some processors structure contracts such that growers receive premiums for fruit based on color and color uniformity. Discoloration due to YSD also reduces lycopene and beta-carotene concentrations in tissue affected by YSD (Darrigues et al., 2008). A reduction in the incidence of YSD could benefit producers, processors, and consumers.

Our objective was to implement a new digital image analysis tool, Color Test (CT), as part of the TA software application (Brewer et al., 2006). The use of flat-bed scanners to acquire data has been reported (Kleeberger and Moser, 2002; Kwack et al., 2005; Shahin and Symons, 2001). However, the software available to analyze color images is not fully automated in these applications and requires many manual adjustments (Shahin and Symons, 2003). TACT is capable of collecting and analyzing color parameters in an efficient, accurate, and high-throughput manner. We evaluated a tomato population for color and color uniformity using TACT and a colorimeter and estimated variance components associated with these parameters. To assess the applicability of TACT to crops other than tomato, we tested the software using images of other fruit and vegetables for which color is an important trait.

## Materials and Methods

### Software implementation

The TA software was previously described by Brewer et al. (2006). Briefly, it was implemented in the programming language C++ using Visual Studio 6.0 (Microsoft Corp., Redmond, WA). The image input/output was made possible via the image processing library Computer Vision and Image Processing 3.7c. TA was designed to run on the Windows operating system, including Vista. The program is free and can be used for academic or private purposes (Van der Knaap, 2008).

### TACT

TACT is designed to collect objective color measurement from JPEG images. We collected images using a flatbed scanner covered with a cardboard box to minimize the effect of shadow and provide a black background (Fig. 1). Images of fruit taken with a digital camera on a black background are also appropriate. Instructions for collecting, importing, and analyzing color from images are available in the TACT manual (Darrigues, 2007).

The TA software automatically recognizes and outlines images of fruit, and the color test module records RGB values of each pixel of the selected object and translates them into average L*, a*, and b* values from the CIELab color space (CIE , 1978). The algorithm implemented in TACT to convert RGB values to L*, a*, and b* values can be adjusted to account for the illuminant (D65 or C) and observer angle (2° or 10°). Converting RGB to L*, a*, and b* was accomplished in three steps according to CIE colorimetry standards (CIE, 2007). First, RGB values were scaled to a perceptually uniform color space [Eq. 1]:

$$\begin{aligned} \text{Var\_R} &= (\{[(\text{R}/255) + 0.055]/1.055\}\text{\^{}}2.4) \times 100 \\ \text{Var\_G} &= (\{[(\text{G}/255) + 0.055]/1.055\}\text{\^{}}2.4) \times 100 \\ \text{Var\_B} &= (\{[(\text{B}/255) + 0.055]/1.055\}\text{\^{}}2.4) \times 100 \end{aligned} \quad [1]$$

Scaled RGB values were then converted to XYZ tristimulus values using the following relationships [Eq. 2]:

$$\begin{aligned} \text{X} &= (\text{Var\_R} \times 0.4124) + (\text{Var\_G} \times 0.3576) + (\text{Var\_B} \times 0.1805) \\ \text{Y} &= (\text{Var\_R} \times 0.2126) + (\text{Var\_G} \times 0.7152) + (\text{Var\_B} \times 0.0722) \\ \text{Z} &= (\text{Var\_R} \times 0.0193) + (\text{Var\_G} \times 0.1192) + (\text{Var\_B} \times 0.9505) \end{aligned} \quad [2]$$

The XYZ values were converted to L*, a*, and b* values using the following relationships [Eq. 3]:

$$\begin{aligned} \text{L}^* &= 116 f(\text{Y}/\text{Yn}) - 16 \\ \text{a}^* &= 500\ [f(\text{X}/\text{Xn}) - f(\text{Y}/\text{Yn})] \\ \text{b}^* &= 200\ [f(\text{Y}/\text{Yn}) - f(\text{Z}/\text{Zn})] \end{aligned} \quad [3]$$

where

$$\begin{aligned} \text{f}(q) &= (q)\text{\^{}}1/3 \quad q > 0.008856 \\ \text{f}(q) &= 7.787q + (16/116) \quad q \leq 0.008856. \end{aligned}$$

Yn, Xn, and Zn are the tristimulus values of the illuminant and observer angle. For illuminant C, observer angle 2°, Xn = 98.04, Yn = 100.0, and Zn = 118.11. For illuminant D65, observer angle 10°, Xn = 94.83, Yn = 100.0, and Zn = 107.38.

The L*a*b × values were then used to calculate chroma as $\sqrt{(a^2+b^2)}$. Hue was calculated as 180/pi × cos[a/$\sqrt{(a^2+b^2)}$] for a* > 0 and as 360 – {(180/pi) × acos[a/$\sqrt{(a^2+b^2)}$]} for a* < 0. The L* coordinate indicates darkness (≈0) to lightness (≈100) of color. Chromaticity coordinates, a* and b*, indicate color directions: +a* is the red direction, –a* is the green direction, +b* is the yellow direction and –b* is the blue direction. Hue is an angular measure from 0 to 360, which represents basic color. Chroma is the saturation or vividness of color. In addition to converting RGB values to L*, a*, and b* and calculating chroma and hue from these components, an algorithm was written for TACT to compute luminosity from the following relationship [Eq. 4]:

$$\text{Luminosity} = (\text{maxCol} + \text{minCol})\ *\ 240.0/(2.0\ *\ 255.0) \quad [4]$$

where maxCol is the highest of the R, G, and B values of an analyzed pixel, and minCol is the lowest value. Luminosity accounts for the variable sensitivity of the human eye to radiation at various wavelengths; it defines brightness. The

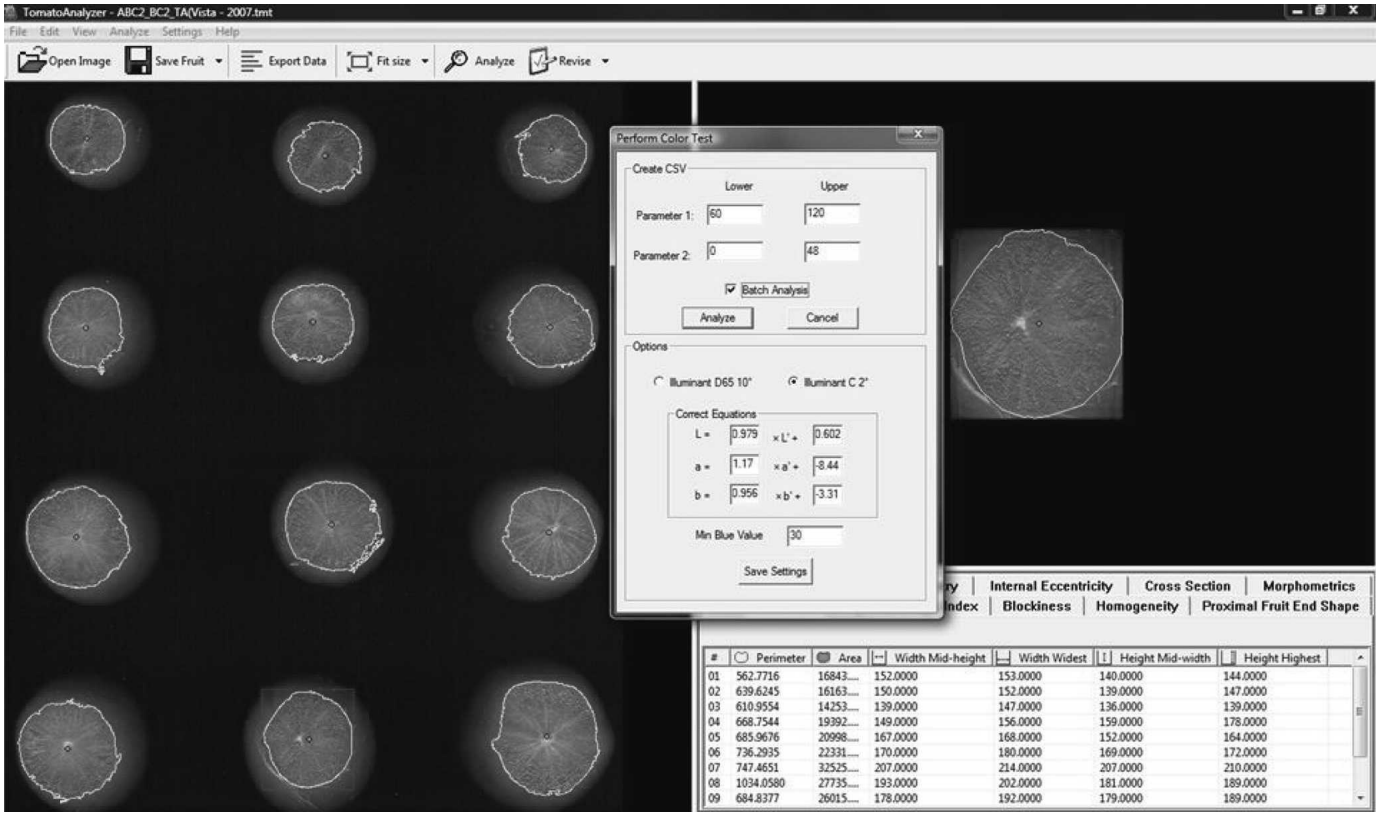


Fig. 1. Tomato Analyzer and its Color Test. The dialog box in the center of the image allows the user to customize the color parameters for analysis (top tier) and to enter the correction values for calibrating the scanner, as well as other options (bottom tier).

output of the color analysis provides averaged values of R, G, B, luminosity, L*, a*, b*, hue, and chroma.

**Customizing TACT.** Several parameters in TACT can be adjusted by the user before analysis. To help differentiate cut surfaces of the fruit from adjacent peel, TACT was designed to allow the user to modify the outline by adjusting the minimum blue value from the RGB color space. In generating the data set for this study, we used a minimum blue value of 30 to define the boundaries. Another option was developed to let the user define two parameters with specific hue ranges of interest. TACT returns the proportion (%) of pixels that fall into the designated hue range. We defined our parameters as percentage of YSD (%YSD), which represents yellow, green-yellow color, and percentage of red (%RED), which corresponds to the desired red color of tomato internal tissue. Our hue ranges were 60 to 120 for %YSD and 0 to 48 for %RED.

**TACT versus colorimeter.** To determine the relationship between colorimeter and TA data, we collected color readings from 247 standard Munsell color plates (X-Rite, Grand Rapids, MI) ranging from 2.5R to 10R (Red), 5Y to 10Y (Yellow), 2.5GY to 10GY (Green-Yellow), 2.5YR to 5YR (Yellow-Red), 5Y (Yellow), 5GY (Green-Yellow), 5G (Green), 5BG (Blue-Green), 5B (Blue), 5PB (Purple-Blue), 5P (Purple), and 5RP (Red-Purple). Absolute color measurements were collected for each standard plate with a colorimeter (CR300; Minolta, Ramsey, NJ). In addition, a JPEG image was collected from scanning each plate with a flatbed scanner (HP Scanjet 3970; Hewlett-Packard, Palo Alto, CA) and was analyzed with the color function of TA. Values for L*, a*, and b* were recorded from each method. Of the 247 standard color plates, 28 were chosen to span a range of colors observed in tomatoes. These plates were custom-made into a 28-patch color checker to be used for scanner calibration (Darrigues, 2007).

We tested the precision of our color measurements with three different scanners: HP Scanjet 3970, HP Scanjet 5300C, and Microtek ScanMaker 6000 (Microtek, Carson, CA). We collected a JPEG image of the 28-patch color checker at 200 dpi with each scanner. Each of the 28 patches was considered an individual object and was analyzed for color. We collected RGB data and converted it to estimates of L*, a*, and b* measurements for each patch using TACT. We also obtained colorimeter L*, a*, and b* data from each patch, and applied linear regression to determine slopes, y-intercepts, and regression coefficients. The TACT dialog box (Fig. 1) allows users to enter correction values for slope and y-intercept as a way to calibrate the device used in collecting images. These correction values, obtained by regression based on color standards, are entered as the inverse of the slope and the negative of the y-intercept for the L*, a*, and b* regressions (Table 1).

**Plant material.** An inbred backcross (IBC) population derived from crosses with *S. lycopersicum* processing cultivars (OH832, OH8245, OH9241, and OH9242) was evaluated for color and color uniformity. The original $F_1$ crosses were OH9242 × OH8245, OH832 × OH8245, and OH9241 × OH8245. Each $F_1$ was backcrossed to the recurrent parent, OH8245, and individual $BC_1$ progeny were again backcrossed to OH8245 to obtain $BC_2$ plants. These $BC_2$ were then selfed to the $BC_2S_4$ generation. Because the recurrent parent was the same and because OH9241, OH9242, and OH832 share a significant portion of their pedigree by descent, the IBC was

Table 1. Correlation coefficients and linear regressions for L*, a*, and b* values for scanning devices.

| Scanner (make and model)[z] | L* | | | a* | | | b* | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r^2$ | Slope[y] | y-intercept | $r^2$ | Slope | y-intercept | $r^2$ | Slope | y-intercept |
| HP Scanjet 3970 | 0.992 | 0.979 | +0.602 | 0.990 | 1.173 | –8.44 | 0.987 | 0.956 | –3.308 |
| HP Scanjet 5300C | 0.991 | 1.020 | –3.65 | 0.985 | 1.391 | –8.96 | 0.977 | 0.979 | –4.381 |
| Microtek Scan Maker | 0.994 | 1.130 | –11.8 | 0.977 | 0.957 | –5.72 | 0.947 | 1.305 | –1.005 |

[z]Hewlett-Packard, Palo Alto, CA; Microtek Inc., Carson, CA.
[y]Regression equations were based on data collected for standard color plates using Tomato Analyzer-Color Test and a CR-300 colorimeter (Minolta, Ramsey, NJ).

considered a single population. Field trials were grown near Fremont, OH, at the Ohio Agricultural Research and Development Center (OARDC) North Central Agricultural Experimental Station in 2004 and 2005. An augmented design was implemented in both years to evaluate each IBC genotype ($n = 179$, $r = 1$) and replicated checks ($n = 4$, $r \geq 5$). Each plot consisted of 20 plants per genotype spaced 30 cm apart, with plots spaced 150 cm apart. All field plots were planted and maintained following conventional practices (Precheur et al., 2004). The plots were harvested when 80% of the fruit were ripe.

**Phenotypic data collection**

Measurements of color were collected using a colorimeter (CR300) as described previously (Sacks and Francis, 2001) and from digital images that were analyzed by TACT. We used a flatbed scanner (HP Scanjet 3970) to scan the cut surface of the proximal end for 12 fruit within each plot and saved a JPEG image. Images were analyzed and color data were collected for two data sets: TA_Unadj, for which no modification was made to the automatic fruit boundaries, and TA_Adj, for which the boundaries were adjusted based on visual inspection to better represent the proximal end of each fruit. Figure 2 provides an example of the boundaries with the two TACT methods used to generate the TA_Unadj and TA_Adj data sets. The batch feature of TACT was used to analyze both data sets, whereby 30 to 50 images were selected per batch and analyzed for color.

Each data set generated from TACT consisted of L*, a*, b*, hue, and chroma values to represent absolute color. For tomato, improved color is characterized by lower L*, b*, and hue values and higher a* values for a more intense, red color. The interpretation of chroma is ambiguous as high chroma due to high b* values represents poor color, whereas high chroma due to high a* values represents good color. In addition, we measured color uniformity defined by the parameters %YSD and %RED.

Two color measurements using the colorimeter were also collected from the same fruit that were scanned for TACT color analysis. The two-point measurements were taken on mature red tissue and any discoloration present on the fruit shoulder. The difference between the two measurements, ΔHue and ΔChroma, provided an estimate of internal fruit color uniformity that is consistent with visual symptoms of YSD.

**Statistical analysis**

All statistical analyses were performed using SAS (version 9.1; SAS Institute, Cary, NC). Color standard data from colorimeter and TA were tested for normal distribution using the UNIVARIATE procedure. To determine the relationship

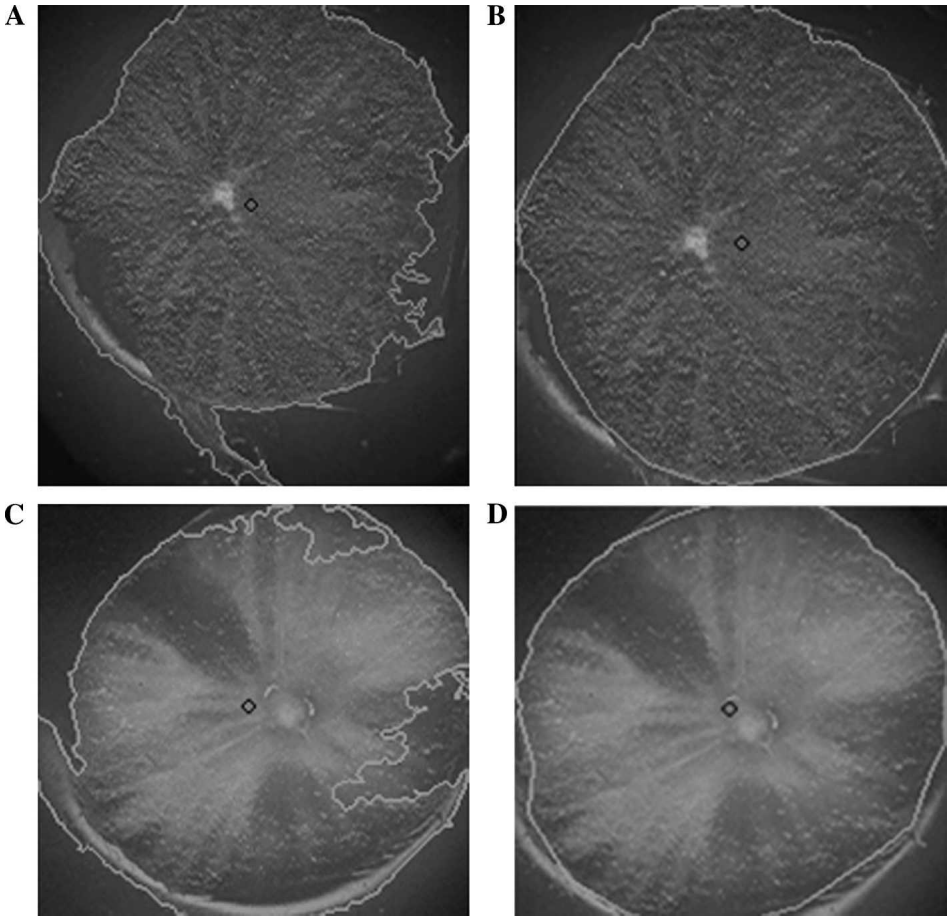


Fig. 2. Representation of the tomato proximal end (shoulder) analyzed for color with Tomato Analyzer-Color Test. **C** and **D** show symptoms of yellow shoulder disorder (YSD), a ripening disorder that affects color uniformity. (**A**) Uniform fruit analyzed using the TA-boundaries defined when minimum Blue = 30 (TA_Unadj method). (**B**) Uniform fruit analyzed with TA-defined boundaries adjusted manually (TA_Adj). (**C**) YSD-affected fruit analyzed with TA_Unadj. (**D**) YSD-affected fruit analyzed with TA_Adj.

between color data generated from the colorimeter and TACT, we used the regression procedure (PROC REG) to test the significance of the regressions.

The estimates of variance and standard errors were obtained using the restricted maximum likelihood (REML) method with the mixed model analysis of variance procedure (PROC MIXED). The model to estimate variance components for the IBC population was:

$$Y_{ijk} = \mu + G_i + T_j + GT_{ij} + \varepsilon_{ijk}$$

where $Y_{ijk}$ was the color trait measured, $\mu$ was the overall mean, $G_i$ was the effect due to the ith genotype, $T_j$ was the effect due to the jth year, $GT_{ij}$ was the effect due to the Genotype × Year interaction, and $\varepsilon_{ijk}$ was the experimental error. The percent total variance was reported to allow direct comparison between data sets. Using the estimates of variance components, broad-sense heritability (H) was determined using the following relationship:

$$H = \sigma^2_G/[(\sigma^2_G) + (\sigma^2_Y/2) + (\sigma^2_{GY}/2) + (\sigma^2_{error}/2)]$$

where $\sigma^2_G$ was the estimate of the genotypic variance, $\sigma^2_G$ was the estimate of the year to year variance, $\sigma^2_{GY}$ was the estimate of the genotype × year interaction variance, and $\sigma^2_{error}$ was the estimate of the error variance. We assumed selection based on IBC line means across 2 years. To test whether methods of collecting data influenced the amount of phenotypic variance that we could partition into genotypic variance, the standard error associated with these estimates was used to perform a mean separation at $\alpha = 0.05$.

## Results

**Correlation between methods.** To determine the precision and accuracy of the color data generated with TACT, we compared it to data collected with a colorimeter. The regressions of L*, a*, and b* values from the colorimeter onto TACT values showed a significant ($P < 0.0001$) linear relationship for all three parameters, with correlation coefficients greater than 0.96 (Fig. 3). Despite the linear relationship, the values between the colorimeter and TACT differed because the slope was not equal to 1 and the y-intercept was not equal to 0. The strong linear relationship suggested that calibration of the scanner used to generate digital images could be accomplished with simple adjustments to the equations used to calculate L*, a*, and b* values.

**Calibration.** Three flatbed scanners were used to assess reproducibility and systematic differences in scanning devices for measuring L*, a*, and b* values from JPEG images (Fig. 4). The correlations between L*, a*, and b* values measured from the colorimeter and TA were high ($r > 0.94$) for all three scanners for each trait. Correlation values among scanners were the highest for L*, lightness, followed by a*, which measures color range from green to red. The lowest correlation values were found for b*, which measures color range from blue to yellow. We observed differences among the three scanners in slope and y-intercept values. For this reason, we implemented an option in the TACT dialog box (Fig. 1) to enter correction values for the slope and y-intercept as a way to calibrate the device used in collecting images.

**Variance partitioning with tact.** To test whether TACT offered advantages over the colorimeter, we evaluated a breeding population for color and color uniformity using both approaches. Variance components for genotype, year, and the interaction genotype × year were estimated to elucidate the proportion of genotypic variance associated with each color parameter (Table 2). Among the three methods, the total phenotypic variation partitioned into genotype and genotype × year interaction ranged from 12% to 30%. The variance partitioned into year ranged from 0% to 9.6%. However, significantly more variance was partitioned into genotype using TACT than the colorimeter for all traits except chroma (Table 2). The proportion of genotypic variance for L*, a*, and b* measured with TACT was 2- to 4-fold greater than with the colorimeter. The ability to partition a greater portion of the phenotypic variation into genetic effects increases the potential for improving a trait by means of genetic manipulation.

In addition to evaluating absolute color (e.g., L*, a*, b*, hue, and chroma), color uniformity was also measured using two user-defined parameters in TACT, defined as %YSD and %RED, and Δ Hue and Δ Chroma from the colorimeter. Variance estimates partitioned ≈10% of the total variation

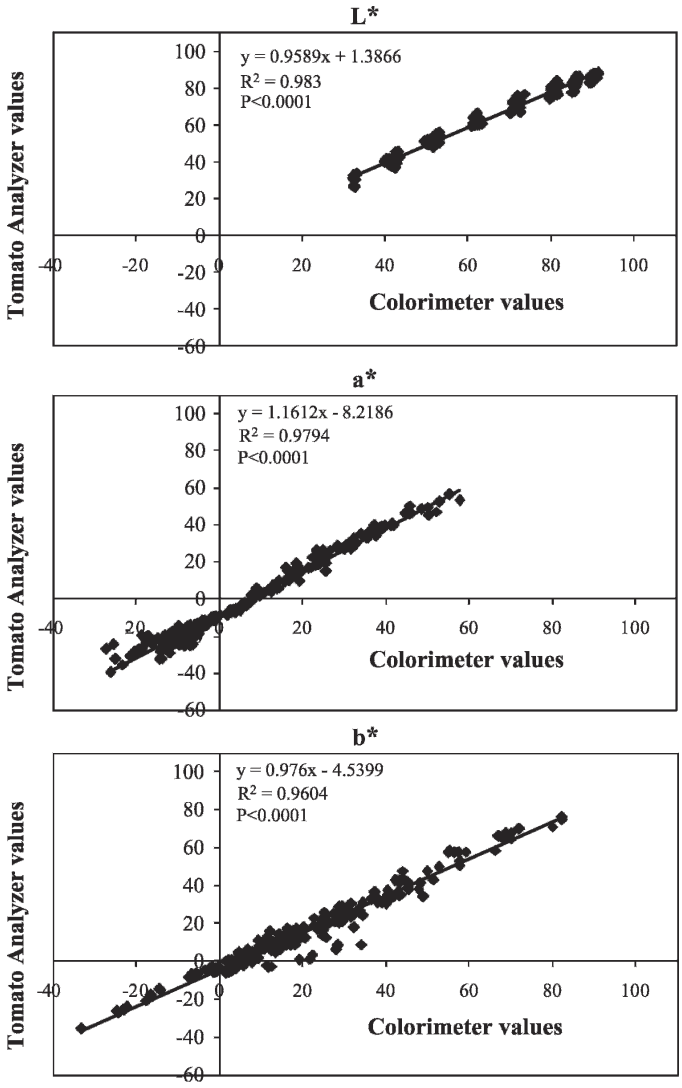


Fig. 3. Correlation between Tomato Analyzer-Color Test and colorimeter values for L*, a*, and b* values of the CIELab color space using data from 247 standard color plates.

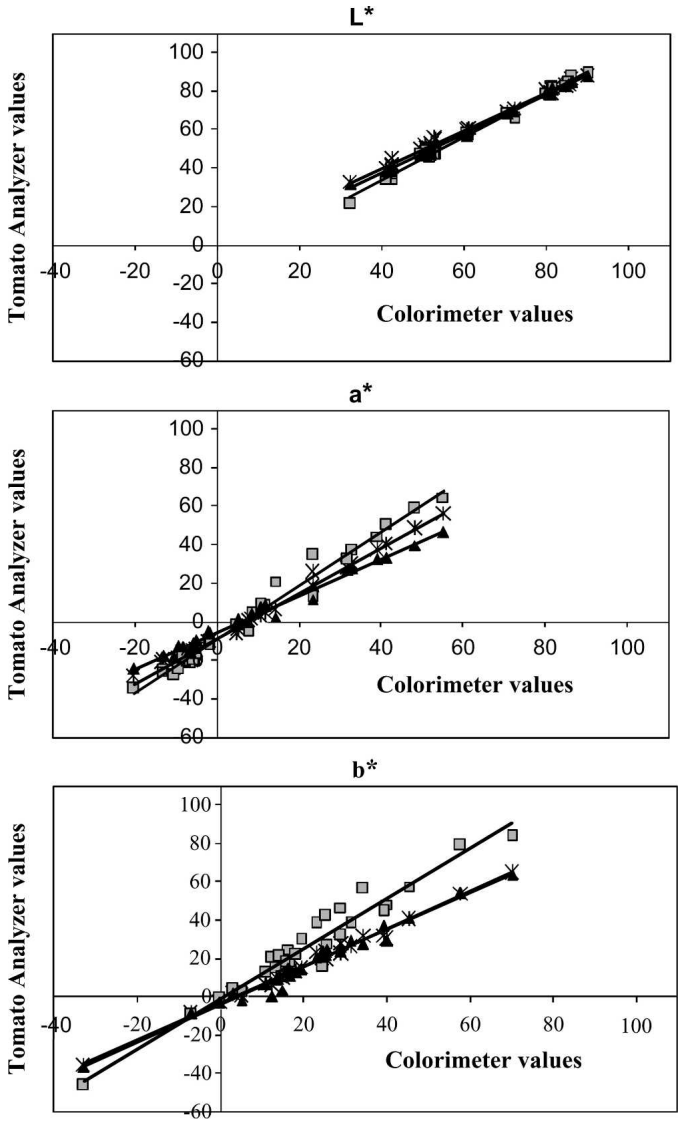


Fig. 4. Regression of L*, a*, and b* values for images from different scanners. Data were obtained from images of standard color plates spanning a range of colors observed in tomato. The scanners used to assess scanning quality were HP ScanJet 3970 (*) (Hewlett-Packard, Palo Alto, CA), HP ScanJet 5300C (▲), and Microtek 6000 (■) (Microtek Inc, Carson, CA). The regression values are summarized in Table 1.

for %YSD and 17% for %RED into the genotypic variance. With the colorimeter, 0% and 4% of the total variance was partitioned into the genotypic variance for Δ Chroma and Δ Hue, respectively. In addition, the error variance was lower for TACT than for the colorimeter for all parameters measured. Overall, 66% to 83% of the total phenotypic variance was in the error term for TACT, compared with 77% to 96% for the colorimeter. In our experimental design, error variance is equal to within-plot variance and complex interactions not controlled for in our sampling. The estimates of genotypic variance correspond to estimates of heritability that ranged from 0.11 to 0.17 for chroma, a*, and %YSD, and 0.211 to 0.275 for L, b*, hue, and %RED when TACT was used. With the colorimeter, H estimates were lower and ranged from 0.081 to 0.097 to L*, a*, and chroma, and 0.113 to 0.152 for b* and hue. With greater heritability estimates, we expect greater gains under selection for genetic improvement by measuring color with TACT.

We generated data sets to test two approaches to define fruit boundaries: TA_Adj and TA_Unadj (Table 2). The TA_Adj data set was compiled from images where boundaries were manually adjusted as needed. The second method, TA_Unadj, was used without making adjustments to the images before color analysis. TA_Unadj is less time-consuming, but is prone to include parts of the fruit, such as peel, deep cracks, or reflected light from the scanner that may bias the color values. However, between the two TA methods, there was no significant difference in the variance partitioned into genetic effects for all color parameters, suggesting that high-throughput analysis with TACT may be possible (Table 2).

**Application of TACT to other crops.** We evaluated TACT with fruit and vegetable crops other than tomato. We tested potato (*Solanum tuberosum* L.), cucumber (*Cucumis sativus* L.), red plum (*Prunus americana* Marsh.), muskmelon (*Cucumis melo* L.), carrot (*Daucus carota* L.), and strawberry (*Fragaria ×ananassa* Duch). These selected crops and cultivars encompassed a range of colors and color uniformity (Fig. 5). TACT recognized the boundaries with precision for each crop, as expected with the contrasting colors between the background and the object (Brewer et al., 2006). The red-skinned potato had the highest L* value, which represents the brightest, closest-to-white tone (Table 3). TACT was able to detect a* values from –8 to 38, with the highest value given to the carrot with its orange tone. The carrot also had the highest b* value, lowest hue, and highest chroma, consistent with the deep orange color.

To test the capability of measuring color uniformity, we defined the hue range of the first parameter as 70 to 120 for the proportion of pixels that fell into yellow-light yellow range. The red plum and muskmelon had 72% to 77% of the pixels falling into that range; the strawberry had 30% (data not shown). For the second parameter, we defined the hue range as 0 to 48 for the proportion of pixels that fell into red-orange range. The carrot had 25% of the pixels falling into that range, whereas the strawberry had only 15% of red-orange tissue. These results demonstrate that TACT can perform color analysis on a broad range of hue values and is not confined to the hue ranges common to the tomato fruit.

## Discussion

We developed a new module in the TA software application to collect objective color measurement based on digital images. This automated tool analyzes each pixel of a selected object and then translates it from RGB to L*, a*, and b* values. Our first objective was to test the accuracy of TACT against a colorimeter and to provide a calibration to account for differences in scanners. Empirical results demonstrated differences between devices due to hardware, software, or nonstandardized RGB values. Digital images collected from different sources can vary in color depending on the resolution, light source, and light quality. Three different scanners were used to scan color standards and test the precision among scanners with TACT. Although the L*, a*, and b* values computed from TACT correlated highly with those of the colorimeter for all three scanners, the slope and y-intercept values varied among scanners. Therefore, an option in TACT was developed to incorporate these values as a correction for L*, a*, and b*. Previously, an attempt to eliminate variability in brightness and color distribution due to scanner differences was reported (Shahin and Symons, 2003). However, the various calibration

Table 2. Proportion of total REML variance estimates for color measurements obtained for fruit from the tomato inbred backcross population using the Tomato Analyzer-Color Test and a colorimeter.

| Method[z] | Variance component[y] | Proportion of REML variance estimates (%)[w] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | L* | a* | b* | Hue | Chroma | Δ Hue | Δ Chroma | % YSD | % Red |
| TA_Adj | Genotype | 12.889 a[x] | 9.839 a | 14.424 a | 11.897 a | 4.923 a | — | — | 9.452 a | 16.636 a |
| | Year | 1.735 | 0.011 | 9.654 | 1.782 | 5.884 | — | — | 0.472 | 2.923 |
| | Genotype × Year | 14.147 | 8.535 | 9.929 | 6.706 | 20.358 | — | — | 6.454 | 10.045 |
| | Error | 71.229 | 81.615 | 65.993 | 79.615 | 68.834 | — | — | 83.621 | 70.396 |
| TA_Unadj | Genotype | 15.120 a | 9.717 a | 15.234 a | 11.705 a | 6.726 a | — | — | 9.659 a | 15.308 a |
| | Year | 0.116 | 0.987 | 5.910 | 3.078 | 1.089 | — | — | 1.325 | 4.285 |
| | Genotype × Year | 12.967 | 8.739 | 10.679 | 7.085 | 20.597 | — | — | 6.338 | 11.399 |
| | Error | 71.797 | 80.557 | 68.176 | 78.132 | 71.589 | — | — | 82.678 | 69.008 |
| Colorimeter | Genotype | 4.229 b | 5.078 b | 8.244 b | 6.002 b | 5.093 a | 3.988 | 0.000 | — | — |
| | Year | 6.102 | 0.000 | 2.803 | 0.381 | 4.166 | 0.465 | 1.285 | — | — |
| | Genotype × Year | 12.978 | 7.269 | 8.766 | 6.610 | 12.892 | 5.452 | 2.198 | — | — |
| | Error | 76.691 | 87.653 | 80.187 | 87.007 | 77.848 | 90.094 | 96.518 | — | — |

[z]Data collected with the Tomato Analyzer (TA)-defined boundaries adjusted manually (TA_Adj), the nonadjusted TA-defined boundaries (minimum blue value = 30; TA_Unadj), and a CR-300 colorimeter (Minolta, Ramsey, NJ).
[y]Data were collected in 2004 and 2005 from plots evaluated in Fremont, OH.
[x]The letter following the estimate of Var(Genotype) represents the statistical grouping for the comparison of each method per trait. Methods for estimating Var(Genotype) in different groupings are significantly different ($\alpha = 0.05$).
[w]REML = restricted maximum likelihood.

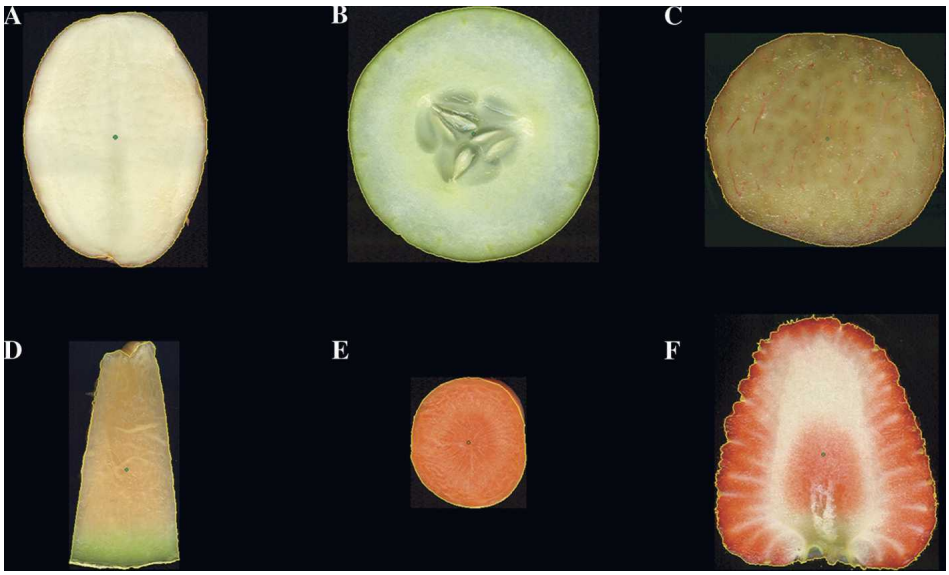


Fig. 5. Images of diverse fruit and vegetables evaluated with Tomato Analyzer-Color Test. (**A**) White-fleshed, red-skinned potato; (**B**) cucumber; (**C**) red plum; (**D**) muskmelon; (**E**) carrot; and (**F**) strawberry.

Table 3. Average values of color parameters obtained from the output of the Tomato Analyzer-Color Test for a cultivar of fruit and vegetables.

| Crop[z] | R | G | B | Luminosity | L* | a* | b* | Hue | Chroma |
|---|---|---|---|---|---|---|---|---|---|
| Potato | 222.2 | 215.7 | 179.9 | 189.2 | 85.82 | –7.925 | 22.74 | 109.0 | 24.48 |
| Cucumber | 183.9 | 197.8 | 149.8 | 163.6 | 77.57 | –16.87 | 26.28 | 123.3 | 31.32 |
| Red plum | 144.7 | 125.7 | 71.63 | 101.8 | 53.17 | –3.037 | 34.27 | 94.62 | 34.76 |
| Muskmelon | 177.7 | 157.8 | 101.4 | 132.0 | 65.53 | –4.173 | 35.03 | 96.66 | 35.90 |
| Carrot | 227.4 | 111.3 | 62.89 | 136.6 | 60.13 | 37.93 | 49.13 | 52.89 | 62.25 |
| Strawberry | 208.2 | 148.9 | 116.6 | 152.8 | 66.99 | 15.38 | 30.24 | 69.64 | 35.71 |

[z]Images of these crops appear in Fig. 5.

techniques exploited did not result in satisfying performance. In contrast, our technique in calibrating TACT was successful, user-friendly, reliable, and malleable.

Our second objective was to assess which technique can better partition observed phenotypic variation for color into genotypic variances. Compared with colorimetric data, the proportion of total phenotypic variance attributed to genotypic variance was significantly improved for all color and color uniformity traits when collecting data with TACT. Estimates of genotypic and phenotypic variances are the basis for determining heritability, which in turn provides an insight into the expected genetic gain and genetic improvement in a breeding program. We show greater line-mean heritability estimates for all traits measured with TACT, with the highest value for %RED. Moreover, the parameters for color uniformity measured with TACT, %YSD, and %RED are more informative in terms of variance estimates than the estimates for ΔHue and ΔChroma measured with the colorimeter. We hypothesize that is the case because color uniformity is better characterized when the entire surface is evaluated rather than the difference between two point measurements. Also, among the methods available with TACT, correcting the boundaries of the tomato fruit cut surface before color analysis (TA_Adj) does not partition significantly more of the genotypic variance from the total phenotypic variance compared with the analysis with unadjusted boundaries (TA_Unadj). The TA_Unadj

is an appropriate method for high-throughput phenotyping as it does not require time-consuming manual adjustments before color analysis.

TACT was designed to be user-friendly with minimum requirements for running it, yet accurate and precise for collecting objective measurements. It facilitates data collection and management, and it requires equipment that is relatively more affordable. Tools for measuring color described in the literature require extensive environmental control, especially for the quality and quantity of light, shadow, and reflection. In contrast, the flatbed scanners we used to generate digital images for TACT color analysis only required a cardboard box as a cover to minimize the effect of shadow. For precision in the analysis of images generated from multiple scanners, we designed a user-friendly option in the dialog box of TACT for calibration relative to color standards. Other methods require a calibration of the system before each image is taken (Lootens et al., 2007) or with minimal calibration feasible (Wang-Pruski, 2006). In our experience, the application of TACT for color analysis from digital images is more accurate, precise, and less expensive than other methods.

Tomato Analyzer, as the name implies, was originally designed to analyze the morphology of tomato fruit. We developed a module for color measurement to expand the array of objective phenotypic analyses implemented. TACT was applied to other fruit and vegetables of various color and color uniformity. Overall, it was able to accurately capture and describe the characteristic color for each crop. Color uniformity was also well characterized for fruit that tend to have nonuniform pigmentation, as in the strawberry (Fig. 5). Its application could go beyond the color analysis of fresh crops. In food science, discoloration after processing or cooking can occur and is an important issue. Wang-Pruski (2006) reports the acquisition of potato tuber images to evaluate after-cooking darkening. Such discoloration could be measured with TACT by defining the specific range of hue values that best represent the undesired discoloration. We show that TACT is a tool that is reliable, precise, amenable, and affordable for digital image analysis of color.

## Literature Cited


Brewer, M.T., L. Lang, K. Fujimura, N. Dujmovic, S. Gray, and E. van der Knaap. 2006. Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other species. Plant Physiol. 141:15–25.

Commission Internationale de l'Éclairage. 1978. Recommendations on uniform color spaces: Color-difference equations, psychometric color terms. CIE Publ. 15 Suppl. 2. Commission Internationale de l'Éclairage, Paris, France.

Commission Internationale de l'Éclairage. 2007. Colorimetry. Part 4: CIE 1976 L*a*b* Colour space. CIE Standard S014–4/E:2007. Commission Internationale de l'Éclairage, Vienna, Austria.

Darrigues, A. 2007. Dissecting variation in tomato fruit color quality through digital phenotyping and genetic mapping. PhD Diss. The Ohio State Univ., Columbus.

Darrigues, A., S.J. Schwartz, and D.M. Francis. 2008. Optimizing sampling of tomato fruit for carotenoid content with application to assessing the impact of ripening disorders. J. Agr. Food Chem. 56:483–487.

Francis, D.M., S.A. Barringer, and R.E. Whitmoyer. 2000. Ultrastructural characterization of yellow shoulder disorder in a uniform ripening tomato genotype. HortScience 35:1114–1117.

Granitto, P.M., H.D. Navone, P.F. Verdes, and H.A. Ceccatto. 2002. Weed seeds identification by machine vision. Computers Electronics Agr. 33:91–103.

Ilic, K., E.A. Kellogg, P. Jaiswal, F. Zapata, P.F. Stevens, L.P. Vincent, S. Avraham, L. Reiser, A. Pujar, M.M. Sachs, N.T. Whitman, S.R. McCouch, M.L. Schaeffer, D.H. Ware, L.D. Stein, and S.Y. Rhee. 2007. The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant. Plant Physiol. 143:587–599.

Kleeberger, K.S. and B.C. Moser. 2002. Flatbed scanners: An alternative tool for gathering horticultural data. HortTechnology 12:444–446.

Kwack, M.S., E.N. Kim, H. Lee, J.W. Kim, S.C. Chun, and K.D. Kim. 2005. Digital image analysis to measure lesion area of cucumber anthracnose by *Colletotrichum orbiculare*. J. Gen. Plant Pathol. 71:418–421.

Leemans, V., H. Magein, and M.F. Destain. 2002. On-line fruit grading according to their external quality using machine vision. Biosystems Eng. 83:397–404.

Li, Q., M. Wang, and W. Gu. 2002. Computer vision based system for apple surface defect detection. Computers Electronics Agr. 36:215–223.

Lootens, P., J. van Waes, and L. Carlier. 2007. Evaluation of the tepal colour of *Begonia* × *tuberhybrida* Voss. for DUS testing using image analysis. Euphytica 155:135–142.

Mendoza, F. and J.M. Aguilera. 2004. Application of image analysis for classification of ripening bananas. J. Food Sci. 69:471–477.

Mueller, L.A., T.H. Solow, N. Taylor, B. Skwarecki, R. Buels, J. Binns, C. Lin, M.H. Wright, R. Ahrens, Y. Wang, E.V. Herbst, E.R. Keyder, N. Menda, D. Zamir, and S.D. Tanksley. 2005. The SOL Genomics Network: A comparative resource for Solanaceae biology and beyond. Plant Physiol. 138:1310–1317.

National Center for Biotechnology Information. 2008. The Entrez Nucleotide database dbEST release 072007. 1 Feb. 2008. <http://www.ncbi.nlm.nih.gov/sites/entrez?db=Nucleotide>.

O'Sullivan, M.G., D.V. Byrne, H. Martens, L.H. Gidskehaug, H.J. Andersen, and M. Martens. 2003. Evaluation of pork color: Prediction of visual sensory quality of meat from instrumental and computer vision methods of color analysis. Meat Sci. 65:909–918.

Picha, D. 2006. Horticultural crop quality characteristics important in international trade. Acta Hort. 712:423–426.

Precheur, R.J., C. Welty, D. Doohan, S.A. Miller, and R.M. Riedel. 2004. Vegetable production guide. Ohio State Univ. Ext. Bul. 672.

Sacks, E.J. and D.M. Francis. 2001. Genetic and environmental variation for tomato flesh color in a population of modern breeding lines. J. Amer. Soc. Hort. Sci. 126:221–226.

Sako, Y., E.E. Regnier, T. Daoust, K. Fujimura, S.K. Harrison, and M.B. McDonald. 2001. Computer image analysis and classification of giant ragweed seeds. Weed Sci. 49:738–745.

Shahin, M.A. and S.J. Symons. 2001. A machine vision system for grading lentils. Can. Biosystems Eng. 43(7):7–14.

Shahin, M.A. and S.J. Symons. 2003. Color calibration of scanners for scanner-independent grain grading. Cereal Chem. 80:285–289.

Tan, J. 2004. Meat quality evaluation by computer vision. J. Food Eng. 61:27–35.

U.S. Department of Agriculture. 2005. Processed products standards and quality certification. 1 Feb. 2008. <http://www.ams.usda.gov/fv/ppb.html>.

Van der Knaap. 2008. Van der Knaap laboratory website. 1 Feb. 2008. <http://www.oardc.ohio-state.edu/vanderknaap/>.

Wang-Pruski, G. 2006. Digital imaging for evaluation of potato after-cooking darkening and its comparison with other methods. Int. J. Food Sci. Technol. 41:885–891.

Yoshioka, Y., R. Ohsawa, H. Iwata, S. Ninomiya, and N. Fukuta. 2006. Quantitative evaluation of petal shape and picotee color pattern in lisianthus by image analysis. J. Amer. Soc. Hort. Sci. 131:261–266.

Zhang, M., J. De Baerdemaeker, and E. Schrevens. 2003. Effects of different varieties and shelf storage conditions of chicory on deteriorative color changes using digital image processing and analysis. Food Res. Intl. 36:669–676.