HortScience 43(3):677–680. 2008.

Fruit Shape Analysis of *Vitis* Using Digital Photography

Andrew P. Wycislo[1,2], John R. Clark, and Douglas E. Karcher
Department of Horticulture, University of Arkansas, 316 Plant Science Building, Fayetteville, AR 72701

Additional index words. grape breeding, small fruits, fruit characterization

***Abstract*. Quantifying fruit shape is challenging, particularly when measurements are made on segregating populations of plants. Objective manual measurements can be performed on small samples of fruit, but this method is difficult and very time-consuming when dealing with larger samples or when shapes are complex or shape variations are slight. Subjective rating scales can also be used, but their effectiveness is questionable when done by multiple raters resulting from varying descriptive standards among individuals. Therefore, a method was developed to analyze digital images containing multiple fruits to characterize fruit shapes. Each segregant of a population of table grapes (*Vitis* spp.) with parents of wide shape variation was photographed and analyzed for shape using SigmaScan® software. The program discriminately selected image pixels representing the fruit and determined the area and perimeter of a grape berry, which were subsequently used to calculate the major:minor axis ratio, shape factor, and compactness values. Computer findings were compared with data from human raters using a simple correlation. When compared with the human ratings, results showed strong correlations of r = 0.941 for major:minor axis ratio, r = –0.804 for shape factor, and r = 0.744 for compactness. This analysis method was a reasonably quick and simple way to quantify grape berry shape, yielding valuable phenotypic data in numerical form. This technology should be useful for shape characterizations in other fruits as well.**

In 1964, Dr. James Moore started the grape breeding program at the University of Arkansas (Clark, 2003). In the early years of the program, Dr. Moore performed some crosses with the cultivar Lady Patricia (Clark, 2003), which was released by the University of Illinois in 1968 (Clark, 1997). This cultivar is unique for its elongated berry shape (Clark, 2003). Dr. Moore believed that the elongated shape was attractive in a table grape, so he continued to intercross selections (J.N. Moore, personal communication). This elongated fruit shape trait has remained in the Arkansas grape breeding program and subsequent crosses have been made to further express this trait. No elongated cultivars have been released, but in the growing area of specialty fruits, these elongated berries could have a promising future in the marketplace.

Although fruit shape is an important characteristic in fruit breeding, few articles exist in the literature that discuss shape as a specific breeding objective. One article discusses flat versus oblong apple (*Malus domestica* L.) shape (Janick et al., 1996). Although it is rare for apples to be taller than wide, it is possible to use the height-to-width ratio of the parents to predict the progeny's ratio as well (Janick et al., 1996). Overall, breeders routinely assess fruit shape; however, those shapes that are odd or undesirable are often discarded or only kept as potential parental material rather than primary selections.

In 2001, a cross was made at the University of Arkansas between the selections A-2315 and A-2659 (J.R. Clark, personal communication). The male parent had large, round to oblong, seedless berries. The female parent was seeded and had very elongated berries. The resulting population exhibited large phenotypic variation, particularly for berry shape. With this population exhibiting continuous variation, it was thought that the genetic markers responsible for regulating fruit shape could be found. The discovery of these markers could assist in future breeding selections by allowing a breeder to determine a selection's ability to produce elongated fruit without waiting several years for a crop.

However, to perform studies of the genetic markers for fruit elongation, the fruit shape must first be quantified. Manual measurements of fruit are impractical as a result of the large number of fruits needing to be tested in a population. Subjective rating scales may work well for one person's analysis but inherently vary with other raters. Large numbers of people would need to rate a population of fruit to create a consistent average. Subjective ratings are also difficult to make because of the continuous variation expressed for this trait. For example, two grapes having only slight differences in shape can be difficult to differentiate using a subjective rating scale or descriptive text. Digital analysis of the fruit would allow for a rapid, more precise method of determining fruit shape measurements. The development of such a system would aid in future genetic studies.

Recently, several researchers have explored the value of having a digital method for analyzing fruit shape. Fruit shape of tomato (*Lycopersicon esculentum* L.) was analyzed by first slicing the tomato in half and placing it on a scanner and then obtaining the measurements using computer software (E. van der Knaap, personal communication). This work, described in Brewer et al. (2006), resulted in the development of a software program called Tomato Analyzer. In their study, they also discussed the need for a better way to analyze shape, specifically height and width. This program used several algorithms to characterize fruit shape attributes, which they also standardized and defined. For example, they defined "fruit shape index" as "the ratio of height to width" and "fruit shape triangle" as "the ratio of the proximal end width to distal end width".

At the University of Arkansas, the turfgrass science program uses a computer program called SigmaScan (SPSS Science, 1999) to digitally analyze turf color and cover. Karcher and Richardson (2003) discussed the need to eliminate human error in turfgrass evaluations. Like with fruit, subjective turf color ratings vary from person to person and do not create a reliable comparison between researchers. Quantifying bare spots in turf objectively is a very time-consuming project requiring manual measurements. SigmaScan's color threshold feature allows one to select a range of colors of interest within an image (Richardson et al., 2001). By selecting a range that encompasses the green turf but not the brown areas around it, the percentage of area covered by turf in a given image can be precisely determined.

Similarly, the program also has the ability to measure the dimensions of an object such as a grape by analyzing an image (SPSS Science, 1999). A pixel is a single point on an image. Because a pixel has a very small area, millions of pixels typically make up one image. The program selects the pixels representing the grape in an image using a color threshold that selects for grape color and not the background of contrasting color. In addition, the program has the ability to differentiate between several objects in one image and can analyze objects in any orientation.

The objective for this study was to develop a simple method to analyze fruit size and shape more quickly using digital images with the computer software program SigmaScan providing data in a detailed, objective form.

Materials and Methods

The population studied was created from a 2001 hybridization of Arkansas table grape selections A-2315 and A-2659. The hybridization, seed collection, stratification, germination, and seedling production were all conducted at the University of Arkansas Fruit Substation, Clarksville, AR. In the

Received for publication 13 Dec. 2007. Accepted for publication 16 Feb. 2008.
[1]Current address: Energy Biosciences Institute, University of Illinois at Urbana–Champaign, 1136 IGB, 1206 W. Gregory Drive, Urbana, IL 61801.
[2]To whom reprint requests should be addressed; e-mail wycislo@uiuc.edu

hybridization, emasculated female parent clusters were protected from contamination by foreign pollen by enclosure in a white, all-weather paper bag. Seedlings from the hybridization were planted at 0.6-m spacing in May 2003. Vines were trained to a single-wire trellis and cane-pruned. Routine culture and pest control procedures were performed on the vines, including drip irrigation as needed, one application of N fertilizer in early spring, and fungicide and insecticide applications near that of a commercial spray program used in Arkansas (Bordelon et al., 2007).

In the summer of 2005, 190 progeny were selected from the population for this study. During the last week of July 2005, fruit samples were collected from 181 of the seedlings and the two parents and placed in separate 3.8-L resealable plastic freezer bags. Fruit could not be collected from nine seedlings, primarily as a result of green June beetle (*Cotinis nitida* L.) insect damage. Representative clusters, usually no more than three per vine, were taken from multicluster vines with the cluster limit being the capacity of the plastic bag. Some vines produced only one cluster, in which case the lone cluster was collected. The grapes were frozen until data were collected.

For data collection, 10 representative berries were selected and removed from each bag and placed on a black surface to be photographed. A Nikon D70S digital SLR camera (Nikon Corp., Tokyo, Japan) was suspended on a copy stand 45 cm above a platform where the berries were placed. The camera was set on the automatic mode and the flash was not used to minimize glare in the resulting images. A platform area ≈15 cm by 30 cm was captured in each image. The same camera positioning and settings were used for each photograph.

One photograph of each set of 10 berries was taken, downloaded to a computer, and appropriately labeled. A SigmaScan macro for image analysis was written in visual basic for applications. The macro analyzed all nonblack objects in each image and determined each object's area, major and minor axis length, and perimeter from each object's pixels. Once the program determined the area and perimeter, three calculations called "major:minor," "shape factor," and "compactness" were applied. Major:minor is the length of the longest axis of an object divided by the shortest axis perpendicular to the long axis. Shape factor is defined as "(4π*area)/perimeter^2" (SPSS Science, 1999). Compactness is defined as "perimeter^2/area" (SPSS Science, 1999).

Data for each berry were output to a Microsoft Excel (Microsoft, Redmond, WA) spreadsheet for each individual item in each image. Every nonblack item on the background one pixel or larger was reported, including small water drops or tiny specks of debris. However, because all anomalies were much smaller (usually less than 500 pixels) than the 10 actual berries (usually greater than 8000 pixels), the noise was easily filtered and the 10 individual data lines representing the berries in each image were moved to a new spreadsheet.

After the digital analysis, a manual rating system was developed. First, all of the berry photographs of the progeny and parents were printed. Next, 14 raters were given instructions to rate the overall elongation of all the grapes in each photo on a scale of 1 to 10, with 1 being round and 10 being very elongated. Then, the ratings were averaged together and correlated to the axis ratio, shape factor, and compactness readings obtained by SigmaScan using the bivariate plot feature in JMP version 6.0 (SAS Institute, 2005). In addition, the manual ratings were individually correlated to the axis ratio, shape factor, and compactness values using SAS version 9.1 (SAS Institute, 2003). Finally, the manual ratings were correlated with each other to determine the consistency among raters also using SAS.

Results and Discussion

The measurements obtained from the SigmaScan program differentiated elongation between the subjects in the population as shape factor decreased with elongation and major:minor axis ratio and compactness both increased. The axis ratio calculations ranged from 1.18 to 2.75, the shape factor calculations ranged from 0.589 to 0.863, and the compactness calculations ranged from 14.56 to 22.85. The distributions of the axis ratio, shape factor, and compactness had skewness readings of 0.783, –0.863, and 1.597, respectively, indicating that the grapes in the population tended toward round rather than elongated (Figs. 1–3).

Visual inspection supported the SigmaScan analysis. The female parent A-2315 (Figs. 4 and 5) is very elongated and had the highest axis ratio at 2.75, the second highest compactness at 21.39, and the second lowest shape factor measurement at 0.589, all indicating greater elongation. The male parent (Fig. 6) is more oblong to round in shape and had a lower compactness of 15.69, a lower axis ratio of 1.38, and a higher shape factor of 0.802. In addition, the progeny with the highest shape factor of 0.863 also had the lowest compactness of 14.56.

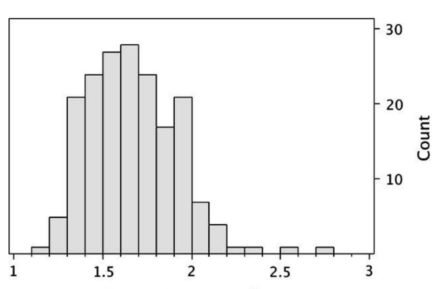

Fig. 1. Distribution of the major : minor axis ratio values. Elongation increases to the right. Female parent A-2315 had a value of 2.75 and male parent A-2659 had a value of 1.39

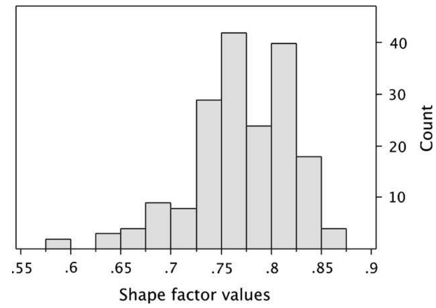

Fig. 2. Distribution of the shape factor values. Elongation increases to the left. Female parent A-2315 had a value of 0.589 and male parent A-2659 had a value of 0.802.

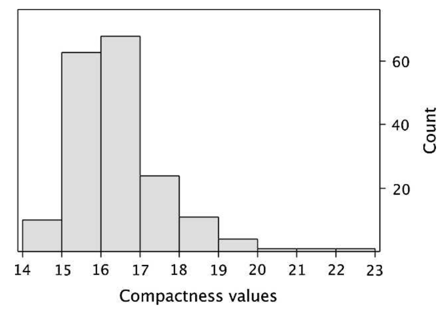

Fig. 3. Distribution of the compactness values. Elongation increases to the right. Female parent A-2315 had a value of 21.39 and male parent A-2659 had a value of 15.69.

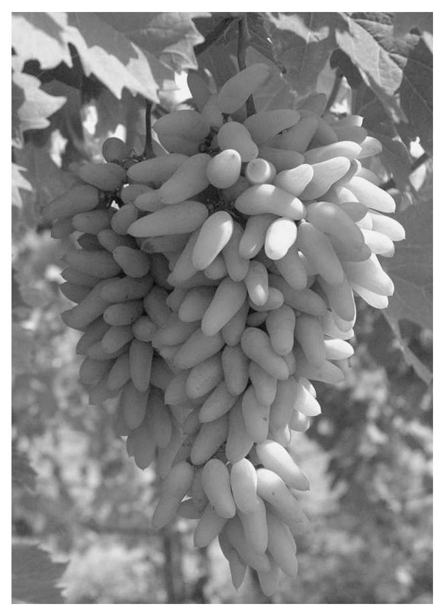

Fig. 4. Elongated female parent A-2315 on the vine.

The manual ratings were compared with the mean axis ratio, compactness, and shape factor measurements using a standard correlation (Figs. 7–9). Results showed strong correlations of 0.941 for the axis ratio, 0.744 for compactness, and –0.804 for shape factor. Shape factor correlation was negative because its equation is inverted in relation to the other two. There were eight outliers, one of which was extreme.

Further statistical analysis showed a very strong correlation of r = 0.986 between shape factor and compactness (Table 1). This indicates that the two calculations performed to determine shape factor and compactness are both good indicators for fruit elongation. However, the correlation between the raters and the axis ratio was much stronger than either shape factor or compactness. The axis ratio is clearly the calculation that is most similar to the manual ratings.

Also, individual correlations between each rater and the axis ratio, shape factor, and compactness were calculated in addition to the mean value correlations described earlier. Again, the correlations were strong. All correlations between raters and digital analysis values were significant at the $P <$ 0.0001 level. However, the correlations were inconsistent across raters (Table 1). These data confirm that the variation among raters can create inconsistent results.

A benefit of this system is that it is nondestructive. The new program ‘‘Tomato Analyzer’’ (Brewer et al., 2006) will likely prove to be valuable to scientists studying fruit shape, but its current protocol suggests cutting each tomato in half before imaging. Besides taking time, this destroys the fruit, making any future study of the fruit difficult. SigmaScan can find the area, perimeter, and all other critical objective measurements of a fruit while it is still intact, preserving it for future study or consumption. Analyzing whole fruits is also useful when studying large populations or small-sized fruits, because manually halving each fruit would slow down the process considerably. In addition, our grapes were frozen after harvest to preserve them until the analysis could be performed. Using this method, the grapes were photographed while they were still frozen and whole, because freezing does not alter the shape.

There are two other major benefits of this system. First, it is very quick and easy to use. One only needs to place the object to be photographed on the surface, capture the image, and move on. Because most digital cameras automatically name the picture files in sequential order, a user would only need to record the order in which objects were photographed to easily match the image with the plant number when the photos are downloaded. Then, the macro used within SigmaScan will automatically analyze all the images in a folder and export the findings to a spreadsheet for further study. Second, the ability of SigmaScan to analyze multiple objects in an image easily allows the user to account for variation. In this study, many of the plants had slight variation among berries within grape clusters. Placing 10 berries in an image and averaging the SigmaScan results together reduced the effect of this variation. More berries could easily be photographed in future studies to reduce this effect even further.

Fig. 5. Berries from female parent A-2315 photographed for digital analysis.

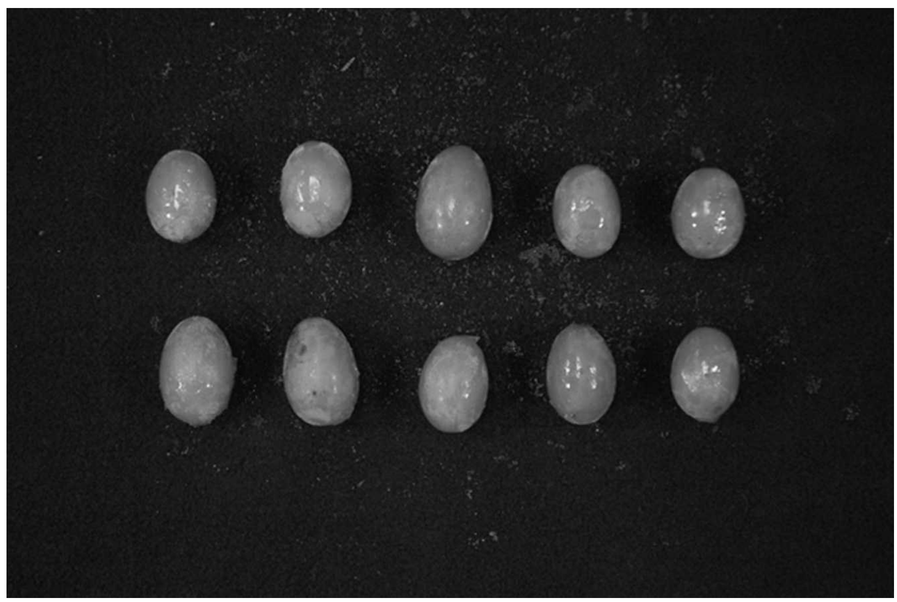

Fig. 6. Berries from male parent A-2659 photographed for digital analysis.

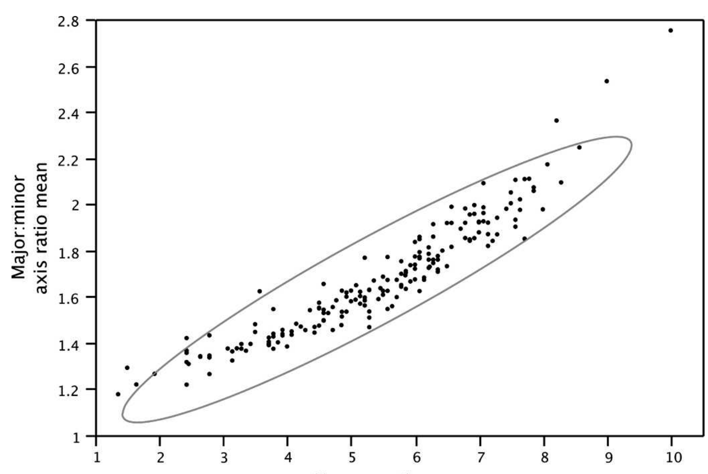

Fig. 7. Correlation plot of human rating mean by major:minor axis ratio mean (r = 0.941). The density ellipse encompasses 95% of the data points.

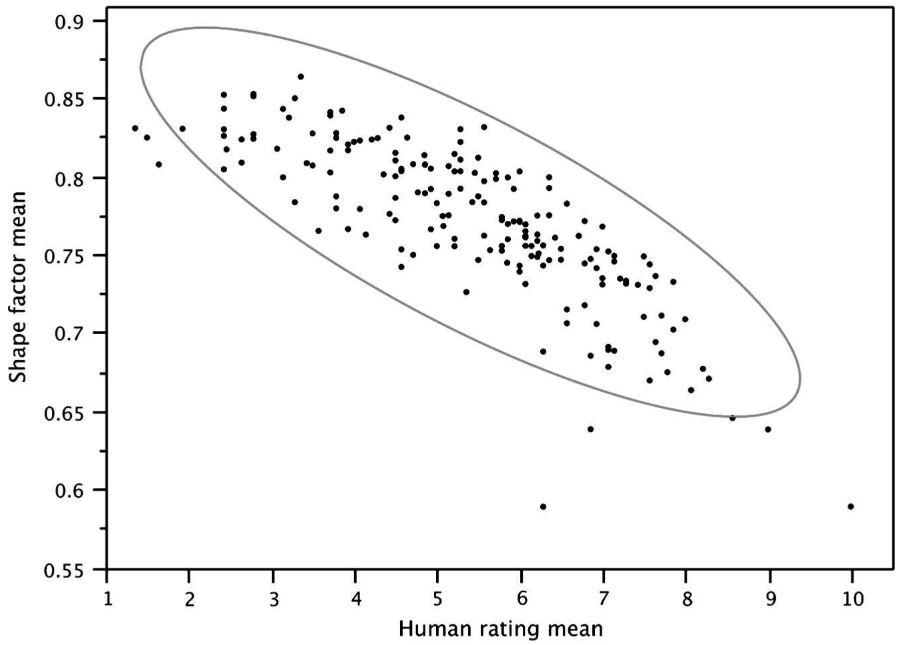

Fig. 8. Correlation plot of human rating mean by shape factor mean (r = –0.804). The density ellipse encompasses 95% of the data points.

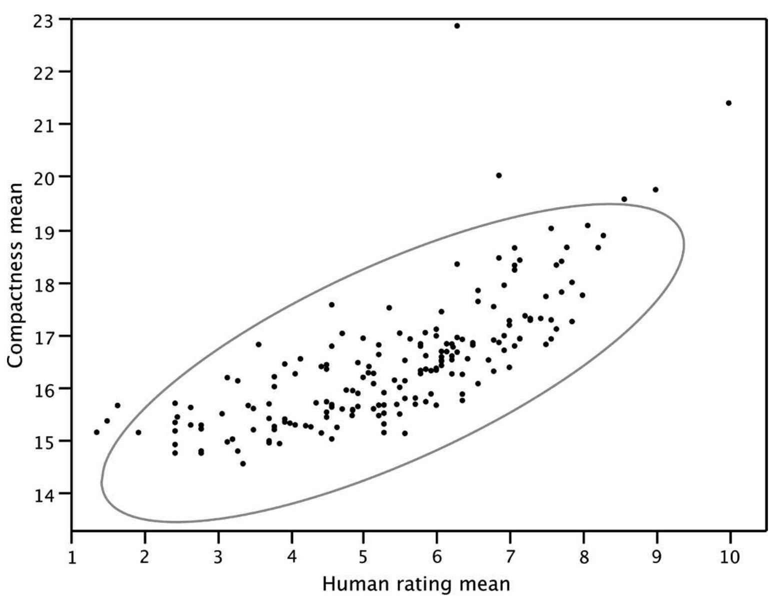

Fig. 9. Correlation plot of human rating mean by compactness mean (r = 0.744). The density ellipse encompasses 95% of the data points.

Table 1. Correlation table of axis ratio (AR), compactness (C), shape factor (SF), and raters (R1–R14).[z]

	AR	C	SF	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14
AR	1	0.79	–0.85	0.86	0.84	0.81	0.92	0.83	0.65	0.87	0.80	0.74	0.88	0.89	0.87	0.85	0.84
C		1	–0.99	0.68	0.68	0.63	0.70	0.71	0.58	0.67	0.57	0.54	0.7	0.72	0.66	0.69	0.68
SF			1	–0.74	–0.74	–0.68	–0.76	–0.75	–0.61	–0.72	–0.63	–0.60	–0.76	–0.78	–0.72	–0.75	–0.72
R1				1	0.83	0.78	0.86	0.78	0.62	0.82	0.82	0.73	0.80	0.87	0.80	0.78	0.81
R2					1	0.74	0.83	0.74	0.59	0.78	0.76	0.67	0.80	0.79	0.80	0.77	0.75
R3						1	0.81	0.74	0.61	0.80	0.83	0.79	0.83	0.77	0.75	0.79	0.69
R4							1	0.81	0.64	0.85	0.80	0.73	0.87	0.85	0.80	0.84	0.78
R5								1	0.61	0.76	0.73	0.70	0.80	0.78	0.76	0.77	0.80
R6									1	0.61	0.56	0.59	0.66	0.65	0.68	0.58	0.63
R7										1	0.77	0.73	0.84	0.82	0.78	0.81	0.74
R8											1	0.79	0.79	0.80	0.75	0.76	0.70
R9												1	0.78	0.78	0.74	0.72	0.69
R10													1	0.83	0.82	0.82	0.79
R11														1	0.81	0.81	0.80
R12															1	0.76	0.82
R13																1	0.75
R14																	1

[z] $P < 0.001$ for the null hypothesis of r = 0.

In conclusion, this digital analysis method is useful for comparisons between members of a grape population for berry shape studies. This study also showed that although manual subjective ratings are useful, a system such as the one described is more repeatable, consistent, and precise. Large numbers of raters would be needed to provide an accurate evaluation for shape. Digital shape analysis outputs data already in objective form, making it useful for studies such as QTL mapping projects. Shape factor and compactness are both good analyses, but the major:minor axis ratio correlated best with the manual ratings. A potential use for this system is by breeders who can use it to help select for grapes with the greatest elongation. Perhaps future studies on shape analysis using SigmaScan will help perfect this method for use on other fruits.

Literature Cited

Bordelon, B., M. Ellis, and R. Foster (eds.). 2007. Midwest commercial small fruit and grape spray guide. 8 Aug. 2006. <http://www.hort.purdue.edu/hort/ext/sfg/>.

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 plant species. Plant Phys. 141:15–25.

Clark, J.R. 1997. Grapes, p. 248–299. Brooks and Olmo register of fruit and nut varieties. 3rd Ed. ASHS Press, Alexandria, VA.

Clark, J.R. 2003. Grape breeding at the University of Arkansas: Approaching forty years of progress. Eighth Intl. Conf. on Grape Genet. Acta Hort. 603:357–360.

Janick, J., J.N. Cummins, S.K. Brown, and M. Hemmat. 1996. Apples, p. 1–77. In: Janick, J. and J.N. Moore (eds.). Fruit breeding. Wiley, New York, NY.

Karcher, D.E. and M.D. Richardson. 2003. Quantifying turfgrass color using digital image analysis. Crop Sci. 43:943–951.

Richardson, M.D., D.E. Karcher, and L.C. Purcell. 2001. Quantifying turfgrass cover using digital image analysis. Crop Sci. 41:1884–1888.

SAS Institute. 2003. The SAS system for Windows. Release 9.1. SAS Inst., Cary, NC.

SAS Institute. 2005. JMP 6.0. SAS Inst., Cary, NC.

SPSS Science. 1999. SigmaScan® version 5.0 user's manual. Chicago, IL.