JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Quantification of Chlorophyll Content and Classification of Nontransgenic and Transgenic Tomato Leaves Using Visible/Near-Infrared Diffuse Reflectance Spectroscopy

LIJUAN XIE, YIBIN YING,* AND TIEJIN YING

College of Biosystems Engineering and Food Science, Zhejiang University, 268 Kaixuan Street, 310029 Hangzhou, People's Republic of China

Visible/near-infrared (vis/NIR) spectroscopy combined with multivariate analysis was used to quantify chlorophyll content in tomato leaves and classify tomato leaves with different genes. In this study, transgenic tomato leaves with antisense *LeETR*1 ($n = 106$) and their parent nontransgenic ones ($n = 102$) were measured in vis/NIR diffuse reflectance mode. Quantification of chlorophyll content was achieved by partial least-squares regression with a cross-validation prediction error equal to 2.87. Partial least-squares discriminant analysis was performed to classify leaves. The results show that differences between transgenic and nontransgenic tomato leaves do exist, and excellent classification can be obtained after optimizing spectral pretreatment. The classification accuracy can reach to 100% using the derivative of spectral data in the full and partial wavenumber range. These results demonstrate that vis/NIR spectroscopy together with chemometrics techniques could be used to quantify chlorophyll content and differentiate tomato leaves with different genes, which offers the benefit of avoiding time-consuming, costly, and laborious chemical and sensory analysis.

KEYWORDS: Visible/near-infrared; partial least-squares; partial least-squares discriminant analysis; tomato leaf

INTRODUCTION

Chlorophyll is one of the most important pigments in higher plant leaves, and it changes throughout the different stages of plant development and between and within species. Therefore, leaf chlorophyll content is a parameter of significant interest from a physiological perspective (*1*).

With the development of plant molecular biology, genetic engineering, genomics, and biochemistry, unprecedented progress has been seen in agriculture, livestock, food quality, nutrition, health, industry, and medicine, and it established the commercial cultivations of transgenic plant varieties which contain new traits including insect resistance, herbicide resistance, delayed fruit ripening, flower color, virus resistance, low amount of nitrates, production of an antifreeze protein, etc. (*2*, *3*). However, there are global issues that arise from the use of genetic modification techniques besides their benefits accruing to consumers, such as the transfer of the introduced genes to wild plants and nontransgenic plants and the indirect effects of the transgenic crops on the environment, modification of the biodiversity of wildlife as a result of changes in the availability of food, and unpredicted harmful changes in their nutritional quality (*2*, *4*). Transgenic products contain an additional trait encoded by an introduced gene, which generally produce an additional protein that confers the trait of interest. Raw material (e.g., grain) and processed products (e.g., food) derived from transgenic products might thus be identified by testing the presence of introduced DNA or by detecting expressed novel proteins encoded by the genetic material (*5*). To monitor and verify the presence and the amount of genetically modified organisms (GMOs) in agricultural crops and in products derived, a demand has been generated for analytical methods capable of detecting, identifying, and quantifying either the DNA introduced or the protein(s) expressed in transgenic plants (*6*, *7*). Ideally an identification technique should be rapid, easy to use, and of low cost (*8*).

Various methodologies that have been employed to analyze and/or detect the presence of GMOs in food products have been developed, such as polymerase chain reaction, enzyme linked immunosorbent assays, biosensor, microarray, chip, electrophoresis, X-ray fluorescence, mass spectrometry, etc. (*9*). These methods based on DNA or protein are versatile, sensitive, qualitative or quantitative, specific, and precise, but they also have some disadvantages including high cost, difficulty of use, need for special equipment, long duration, and so on (*5*, *10*). In many situations, a test will be required to detect the presence of GMOs in commodities or food. This requirement for qualitative or quantitative analysis will impact on the most appropriate testing method for that application. Near-infrared (NIR) spectroscopy, allied to multivariate calibration techniques, has gained wide acceptance in different fields (*11*). The advantage of NIR spectroscopy versus other previously used techniques lies in its nondestructive, simple, fast nature, which

* To whom correspondence should be addressed. Tel.: 0086-571-86971140. Fax: 0086-571-86971885. E-mail: ybying@zju.edu.cn.

10.1021/jf063664m CCC: $37.00
Published on Web 05/16/2007

makes this technology ideally suited for on-line process monitoring and quality control (*12*). Though NIR spectrometers are not precise enough to detect compounds at the DNA concentration level (parts per trillion), spectral differences caused by larger structural changes (if any) accompanying the modification might be measurable.

NIR spectroscopy has increasingly been adopted as an analytical tool in various fields. One of its most common applications combined with chemometrics methods for classification, such as soft independent modeling of class analogy (*13*, *14*), principal component analysis (*15*), partial least-squares discriminant analysis (PLSDA) (*16*), artificial neural networks (ANNs) (*17*), linear discriminant analysis (*18*), locally weighted regression (LWR) (*19*), and so on, has been used to discriminate samples belonging to one of several distinct groups based on their spectral properties (*18*).

Recently, this technique has been used to distinguish transgenic products from conventional ones. Roussel et al. (*19*) detected and segregated Roundup Ready soybeans from conventional soybeans using partial least-squares (PLS), LWR, and ANN models by NIR spectroscopy. A total of 93% accurate classification was obtained using a database of approximately 8000 samples with the LWR method. Rui et al. (*20*) applied back-propagation algorithm to discriminate transgenic corns and their parents by continuous wave of NIR diffuse reflectance spectroscopy range of 4000−12 000 cm^{-1}.

The objective of this study is to examine the feasibility of using the visible/near-infrared (vis/NIR) diffuse reflectance spectroscopic techniques to quantify chlorophyll content and distinguish transgenic tomato leaves with antisense *LeETR*1 from nontransgenic tomato leaves. The specific goals were to (1) investigate the potential of vis/NIR spectroscopy in quantification of chlorophyll concentration in tomato leaves and discriminate the differences depending on this composition, (2) detect the presence of transgenic tomato leaves, using PLSDA for classifying vis/NIR leaf spectra into two groups, nontransgenic tomato leaves and transgenic ones, and (3) evaluate the classification accuracy of PLSDA models.

MATERIALS AND METHODS

Samples. Transgenic tomato plants were produced by using standard *Agrobacterium tumefaciens* mediated plant transformation methods. Polymerase chain reaction and southern-blot analysis were used to check the integrity and number of copies of the introduced genes (*21*). Pure transgenic plants with antisense *LeETR*1 and nontransgenic plants were grown under standard greenhouse conditions. One hundred plants of each variety were grown, and at least one leaf was collected from one plant. A total of 106 transgenic leaves and 102 nontransgenic ones with similar sizes were quickly picked for vis/NIR diffuse reflectance measurements.

Chlorophyll Content Measurements. A chlorophyll meter measures transmission of red light at a wavelength where chlorophyll absorbs light and transmission of infrared light at a wavelength where no absorption occurs. On the basis of these two transmission values, the instrument calculates a SPAD value that is quite well correlated with chlorophyll content (*22*). In this study, a hand-held chlorophyll meter (SPAD-502 chlorophyll meter, Minolta Camera Co., Ltd., Japan) was used to measure leaf chlorophyll content. Chlorophyll content is different over one leaf, but at a small area, we suppose the distribution of chlorophyll content is same. The chlorophyll meter can get chlorophyll content at a small area where the vis/NIR spectra were collected. Every leaf was measured three times, and the average value was denoted as the chlorophyll content of that leaf. **Table 1** shows the descriptive statistics for the chlorophyll concentrations analyzed in both tomato leaf varieties. By comparison of the chlorophyll concentrations of the samples, a statistically significant difference in the values was found using a Student *F* test ($\rho < 0.05$).

Table 1. Range, Mean, Standard Deviation (SD), and Significance of Differences in Chlorophyll for Transgenic and Nontransgenic Tomato Leaves

	range	mean	SD	significance of difference[a]
transgenic tomato leaves	38.5–61.4	49.61	5.97	*
nontransgenic tomato leaves	45.7–65.4	53.68	3.81	*

[a] Asterisk indicates significance, $\rho < 0.05$.

Spectroscopic Measurements. The vis/NIR diffuse reflectance spectra were collected with a FT-NIR spectrometer system (Thermo Electron Corp., Madison, WI) fitted with an optic fiber cable, a cooled Si detector (9000−15000 cm^{-1}), and a 50 W quartz halogen light source. In the head of the bifurcated cable, both light source beams and receptor beams were enclosed in the fiber probe randomly.

The spectrometer was connected via a PCI card to a personal computer, and specific software OMINIC 6.1a (Thermo Electron Corp., Madison, WI) was available to modify spectrometer setup and store acquired spectra. The mirror velocity was 0.9494 cm s^{-1}, and the resolution was 8 cm^{-1} in this work. Each spectrum was the average of 32 successive scans.

Leaves were placed directly upon the fiber probe by hand with the same orientation, the adaxial epidermis of leaf was contacted with probe's surface closely, and the reflectance spectrum of each leaf was obtained. On each leaf, a diffuse reflectance spectrum was measured randomly three times, and only the averaged spectrum of the three spectra was used for analysis and stored as the logarithm of the reciprocal of the reflected (R) energy ($\log(1/R)$). Before fruit spectra acquisition, a reference spectrum was collected from a standard white Teflon cylinder. The vis/NIR measurements were performed at 20 °C in a dark room only in 1 day, and spectra collections for transgenic samples and nontransgenic ones were alternate.

Spectral Data Pretreatment. NIR spectra are affected by both the concentration of the chemical constituents and the physical properties of the analyzed product, and the latter properties account for the majority of the variance among spectra (*23*). It means that physical properties, such as particle size and distribution affect spectra strongly, while chemical composition is considered as small. It is necessary to perform mathematical pretreatments to reduce the effects of scatter (*24*) and enhance the contribution of the chemical composition. In this study, the pretreatments used for the calibrations were multiplicative scattering correction (MSC) and the first and second derivative (*25*).

Data Analysis. Chemometrics analysis was performed using the commercial software package TQ Analyst v6.2.1 (Thermo Nicolet Corporation, Madison, WI).

The chlorophyll of tomato leaf quantification models were developed using partial least-squares regression. Leave-one-out cross-validation was performed to estimate the prediction error and detect any outliers. In this validation, all samples except one are used to construct a PLS model, and then the model is used to predict the remained sample. After that, a second sample is left out from all samples, and a newly constructed model is used to predict the second sample. This procedure is repeated until each sample is left out and predicted by a model once. In this step, spectrum outlier diagnostic, which finds the spectra of the standards which are most unlike the spectra of the other standards and uses either the Dixon or the Chauvenet test for outliers to determine whether the difference is significant, was run. In this diagnostic, the spectral and concentration information for each component and standard are used to determine Mahalanobis distance values. The justification criterion is similar to that reported by Liu (*26*).

Discriminant studies to classify tomato leaves with different genes were performed using PLSDA. PLSDA is a PLS application for the optimum separation of classes, and each sample was assigned a dummy variable 1 or 2 as a reference value, an arbitrary number which indicates whether the sample belongs to a particular group or not (*14*). In this case, samples of transgenic tomato leaf were assigned a numeric value of 1, and those of nontransgenic leaf were assigned 2. The PLSDA model was then developed by assigning the reference value (dummy

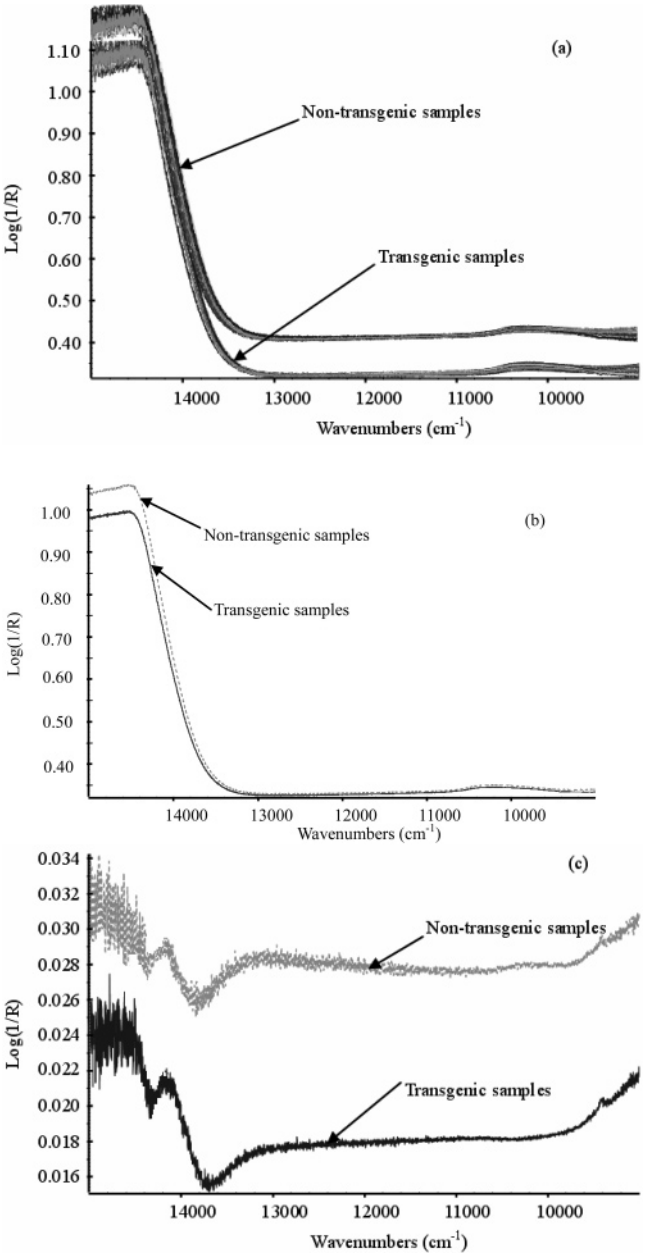

Figure 1. (**a**) Raw spectra, (**b**) average spectra, and (**c**) standard deviation of raw spectra of the transgenic and nontransgenic tomato leaves.

variable) for each sample. A sample was considered to be correctly categorized if the predicted value lay on the same side of the midpoint of the assigned values (*27*). The transgenic sample was classified correctly if the value was between 0.5 and 1.5, else the sample was classified wrong. It was a nontransgenic sample if the value was between 1.5 and 2.5. The criteria for the cutoff selected were similar to those reported by Cozzolino and Andre (*15*, *16*). In this step, 150 samples (75 transgenic samples and the same number of nontransgenic ones) were used for calibration, and the remaining 58 ones for validation. The samples for calibration and validation sets were chosen randomly.

Wavelength selection not only enhances the stability of the model resulting from the collinearity in multivariate spectra but also helps to interpret the relationship between the model and the sample compositions. It is also crucial to select an appropriate portion to improve the performance of a calibration model (*28*). The wavenumber cutoff beyond 14881 cm^{-1} and under 9091 cm^{-1} was selected due to low signal-to-noise ratio which is typical for devices equipped with fiber optics. Therefore, only the spectral region ranging from 9091 to 14881 cm^{-1} was used. In this research, according to our FT-NIR spectrometer and TQ Analyst software, which measured the intensities of all the

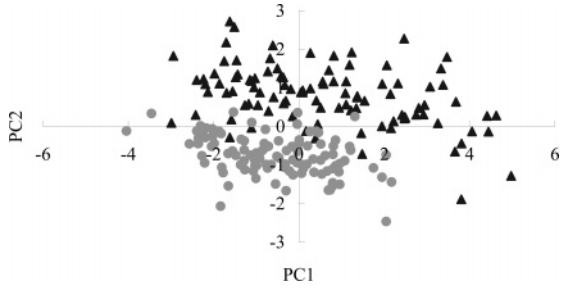

Figure 2. First two PC score plot of tomato leaves using raw NIR spectra. Circles represent transgenic samples, and triangles represent nontransgenic samples.

data points using the algorithm specified in the region type in the software and suggested spectrum range to develop robust models, the spectra were divided into two regions. Region one ranged from 9091 to 12 500 cm^{-1} and region two from 12 500 to 14 881 cm^{-1}. Each region has at least one peak and one shoulder. For separate regions or combinations of these regions, PLS and PLSDA models with different loadings were investigated. The optimal spectral range and model size were then selected and determined by the lowest value of the predicted residual error sum of squares. It is expected to have ideal models with the lower root-mean-square error of calibration (RMSEC), root-mean-square error of cross-validation (RMSECV), and root-mean-square error of prediction as well as the higher correlation coefficient *r*.

All of these different procedures were evaluated to find the spectral region and the data pretreatment method that give the best prediction and classification.

RESULTS AND DISCUSSION

Spectral Analysis. Spectra of tomato leaves are shown in **Figure 1a**. The transgenic tomato leaf spectrum is offset by some units for clarity. It is readily apparent that no significant difference that can be observed by the naked eye exists between them. As **Table 1** shows, one difference between these varieties is chlorophyll content. In the case of transgenic tomato leaves, the mean value of chlorophyll is 49.61; however, for nontransgenic tomato leaves, the mean value is 53.68, a little higher than that of transgenic ones. **Figure 1b** shows the average spectra of the transgenic and nontransgenic tomato leaf samples analyzed. Obvious differences were found from a visual observation of the two spectra. The spectra reveal peaks both at 10 310 cm^{-1} representing stretch vibrations associated with the O−H group in water and at 14 550 cm^{-1} related to pigment absorption where a large gap is indicated between these lines. **Figure 1b** also indicates that the absorption at every wavelength for the averaged spectrum of transgenic tomato leaves is remarkably lower than that of nontransgenic tomato leaves. As illustrated, absorption at 14550 cm^{-1} is the highest. Even though chlorophyll content might partially explain the observed spectral difference between types of leaf, the difference might arise from other facts. (1) Besides chlorophyll content, there are other different compositions, such as water. (2) There may be a different percentage of ethylene resulting from the introduced antisense *LeETR*1. (3) Since ethylene acts to promote the transcription and translation of numerous ripening-related genes, including those involved in cell wall breakdown and carotenoid biosynthesis (*29*), some ingredients such as carotenoid may be changed. The high variations in the raw spectra can be seen in **Figure 1c**, which shows the standard deviation of spectral data at 6223 points and were around 13 080, 13 855 and 14 250 cm^{-1}. As mentioned earlier, due to low signal-to-noise ratio beyond 14881 cm^{-1} and under 9091 cm^{-1}, the standard deviation there was very high.

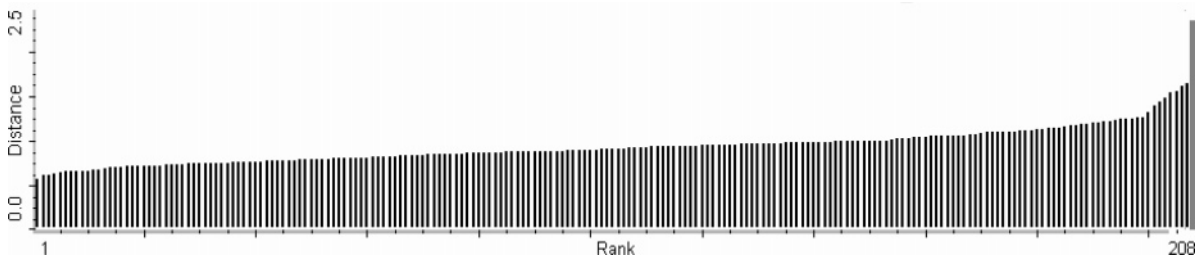

Figure 3. Distribution of Mahalanobis distance values of 208 tomato leaf samples.

Table 2. Statistics for Leaf Samples Using PLS Regression Models (Raw and MSC) ($n = 208$)

data type	wavelength range (nm)	r	RMSEC	RMSECV	loadings	outliers
raw	9091–14881	0.95208	1.65	2.83	6	1
	9091–12500	0.67212	4.01	4.68	5	1
	12500–14881	**0.95567**	**1.59**	**2.87**	**5**	**1**
MSC	9091–14884	0.95414	1.62	2.81	4	1
	9091–12500	0.86462	2.82	2.97	4	1
	12500–14881	0.94534	1.76	2.93	3	0

Overall, transgenic tomato leaves and nontransgenic tomato leaves are very complex, which results in vis/NIR diffuse reflectance spectra that are highly overlapped; however, unique spectral differences between two samples can provide important information.

To investigate the feasibility of discrimination between the samples of the two varieties, the first and second score plots were shown (**Figure 2**), derived from the raw spectra of the samples. PC1 accounts for 70.14% of the variation in the spectra, and PC2 explains 24.71% of the variation in the spectra. From this figure, we can find that the samples are divided roughly into two groups. There is no clear boundary, and many points overlap each other. It can also be seen that nontransgenic samples are mostly located on the top half of the plot while transgenic ones are observed on the reverse half in a great measure, which indicated that nontransgenic samples have positive scores in the second component and transgenic ones have negative ones.

Quantification Analysis. A total of six prediction models were developed for the quantification of chlorophyll content in transgenic and nontransgenic tomato leaf samples. These involved the use of two forms of spectral data—raw and MSC. The summary results of this work are presented in **Table 2**. The PLS models are good for predicting chlorophyll concentrations, and their prediction accuracy can be improved by using the MSC process with the exception of the range of 12 500−14 881 cm^{-1}. For raw spectral data, the wavenumber range 12 500−14 881 cm^{-1} only required three loadings and also had a relatively low RMSEC of 1.76, r equal to 0.94534. We also can find that models using the whole spectra with or without suitable data handing may give a relatively robust prediction, but for the models using 9091−12 500 cm^{-1} wave band, the results were not so satisfactory. It might arise from the fact that pigment, such as chlorophyll in leaf, makes a contribution to the vis/NIR models of tomato leaves. Overall, the most accurate model may be that which involved raw data in the wavenumber range 12 500−14 881 cm^{-1}, using five PLS loadings and producing RMSEC and RMSECV equal to 1.59 and 2.87, respectively. In this model, only one sample was removed. **Figure 3** shows the Mahalanobis distance values of the 208 original leaf samples ranked from smallest to largest. As shown in figure, when a threshold of 1.640 is selected, one sample (Mahalanobis distance = 2.341) is rejected for its large Mahalanobis distance.

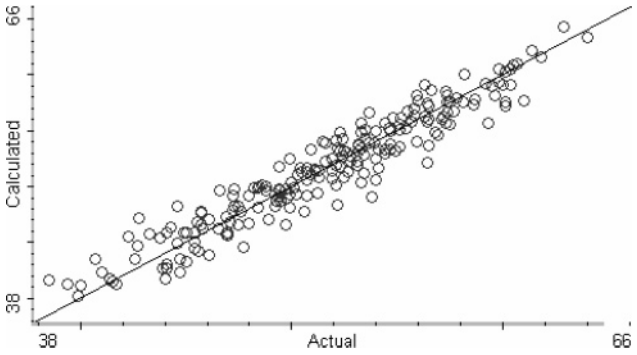

Figure 4. Correlation statistics between the measured values and calculated values of chlorophyll content in transgenic and nontransgenic tomato leaves.

Table 3. Statistics for Leaf Samples Using PLSDA Models (Raw, First, and Second Derivative) ($n = 208$)

data type	wavelength range (nm)	r	RMSEC	RMSECV	% of samples correctly classified in calibration	loadings
raw	9091–14881	0.71996	0.347	0.3473	83.7	1
	9091–12500	0.57663	0.408	0.4083	77.9	1
	12500–14881	0.76404	0.323	0.3234	87.0	1
first derivative	9091–14881	0.95543	0.148	0.1483	100	1
	9091–12500	0.92429	0.191	0.1914	100	1
	12500–14881	0.94528	0.163	0.1635	100	1
second derivative	**9091–14881**	**0.95915**	**0.141**	**0.1414**	**100**	**1**
	9091–12500	0.90161	0.216	0.2165	98.1	1
	12500–14881	0.94883	0.158	0.1582	100	1

Figure 4 shows the correlation between the values determined by the reference analysis and the values predicted by the vis/NIR spectroscopy technique on the whole sample set of tomato leaves. The diagonal line represents ideal results (actual = calculated value), and so the closer the points are to the line, the better is the model.

Classification. Table 3 contains the statistics for tomato leaf samples using PLSDA regression models developed either on the raw spectra or after first and second derivative spectra using various wavenumber regions. In agreement with Cozzolino, the optimal combination of spectral regions varied among regression methods (*30*). It can be found that the full spectral region using the first and second derivative data gave better calibration statistics, but for raw data, the wavenumber region of 12500−14881 cm^{-1} gave a relatively good result. The derivative process can increase classification, and the percent correct classifications for transgenic and nontransgenic tomatoes ranged from 83.0%

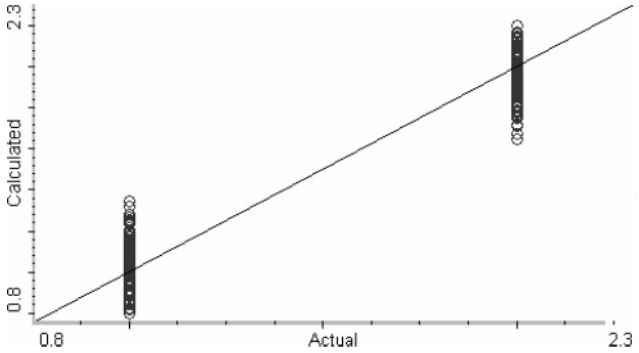

Figure 5. Prediction of tomato leaf varieties using PLSDA regression (9091–14881 cm^{-1}).

Table 4. PLSDA Classification Results for Leaf Samples Using NIR Spectra (Raw, First, and Second Derivative) ($n = 208$)

		correct classification			
		transgenic samples		nontransgenic samples	
data type	wavelength range (nm)	calibration set	validation set	calibration set	validation set
raw	9091–14881	68	26	63	17
	9091–12500	65	23	59	15
	12500–14881	70	28	65	18
first derivative	9091–14881	75	31	75	27
	9091–12500	75	31	75	27
	12500–14881	75	31	75	27
second derivative	9091–14881	75	31	72	27
	9091–12500	74	31	73	26
	12500–14881	75	31	75	27

to 100% and 72.5% to 100%, respectively, depending on the wavenumber range and data type used (**Table 4**). Table 4 shows that, in the calibration set, the lowest classifications for transgenic and nontransgenic tomato leaves were 86.7% and 78.7%, and in the validation set they were 74.2% and 55.6%, respectively. The use of raw spectral data in the range of 9091−12 500 cm^{-1} produced a very high level of error classification. A total of 28 nontransgenic tomato leaves and 18 transgenic ones were misclassified with raw spectra in the wavenumber region of 9091−12 500 cm^{-1} (**Table 4**). The wrong classified samples might be classified to other groups or could not be identified. Overall, the best model derived was that produced by the second derivative treatment of spectral data over the entire wavelength range; this gave a 100% correct classification with $r = 0.95915$, RMSEC value of 0.141, RMSECV value of 0.1414, and a factor in the model. It is a stable and robust model for its low and similar RMSEC, RMSECV, and high r. Models based on first and second derivatives of spectra data in the range 12 500−14 881 cm^{-1} produced models of relatively high accuracy. These two models have the same high correct classification, both 100% correct classifications for transgenic and nontransgenic tomato leaves. This phenomenon could be attributed to pigment absorptions.

Figure 5 shows the vis/NIR predictions of tomato leaf varieties using the PLSDA model with the spectra after second derivative treatment in the full spectral region. Nontransgenic tomato leaves with predicted values ranged form 1.5 to 2.5 and transgenic samples from 0.5 to 1.5 were all considered to be correctly classified by the model. On the basis of the vibrational responses of chemical bonds to vis/NIR radiation, the model can discriminate or identify varieties. It is probable that the higher the variability between sample types in those chemical entities, which respond in these regions of the spectrum, the better the accuracy of the model (*15*). It suggests that PLSDA models with spectra after derivative treatment obtained enough information for discriminating the samples because of their difference in the chemical components. However, relative to the nature of the method used, it was not possible to define a single compound or a group of compounds that explain the differences observed between the two varieties studied.

Conclusions. This study shows the potential of vis/NIR spectroscopy to quantify chlorophyll content and classify transgenic and nontransgenic tomato leaves. Calibration models relating spectral characteristics of samples from tomato leaves with chemical composition, regarding chlorophyll, were successfully built using relatively simple models and variable selection techniques. With regard to chlorophyll content in leaves, the PLS regression model produced low RMSEC and RMSECV equal to 1.59 and 2.87, respectively. The study also indicates that differences between transgenic and nontransgenic tomato leaves do exist and groups are apparent. Visible/near-infrared spectroscopy, combined with multivariate analysis after the appropriate spectral data pretreatment, has been proven to be a very powerful tool for judgment of a relative pattern among objects that have very similar properties. The greatest classification accuracy was achieved using the second derivative of spectral data in the full wavelength range, with accuracy up to 100%. The methods got comprehensive and complementary information to distinguish tomato leaves with different genes. Classification methods offer the benefit of avoiding time-consuming recalibration work for each sample and costly and laborious chemical and sensory analysis. Further studies are needed to use chemical parameters expressing chemical quality of tomato to develop valuable and robust models to discriminate tomato varieties or blends.

LITERATURE CITED

(1) Richardson, A. D.; Duigan, S. P.; Berlyn, G. P. An evaluation of noninvasive method to estimate foliar chlorophyll content. *New Phytol.* **2002**, *153* (1), 185−194.

(2) Kauser, A. M. *Advancement in Biotechnology*; NIBGE-FAO Workshop on GMO detection, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan, 2004, pp 8−11.

(3) Gachet, E.; Martin, G. G.; Vigneau, F.; Meyer, G. Detection of genetically modified organisms (GMOs) by PCR: a brief review of methodologies available. *Trends Food Sci. Technol.* **1999**, *9* (11), 380−388.

(4) Perry, J. N. *GM Crops and the Environment*; Proc. Cambridge Soc. Application, 2002.

(5) Ahmed, F. E. Detection of genetically modified organisms in foods. *Trends Biotechnol.* **2002**, *20*, 215−223.

(6) Bonfini, L.; Heinze P.; Kay S.; and Van Den Eede, G. *Review of GMO Detection and Quantification Techniques*, final version; Food Products and Consumer Goods Unit, I-21020 Ispra, Italy, 2001.

(7) Anklam, E.; Gadani, F.; Heinze, P.; Pijnenburg, H.; Guy, V. D. E. Analytical methods for detection and determination of genetically modified organisms in aricultural crops and plant-derived food products. *Eur. Food Res. Technol.* **2002**, *214*, 3−26.

(8) Belton, P. S.; Kemsley, E. K.; McCann, M. C.; Ttofis, S.; Wilson, R. H.; Delgadillo, I. The identification of vegetable matter using Fourier transform infrared spectroscopy. *Food Chem.* **1995**, *54* (4), 437−441.

(9) Zafar, Y.; Asif, M. *Capacity Building in Biosafety of Genetically Modified Crops: GMOs (Genetically Modified Organisms) Detection*; National Institute for Biotechnology and Genetic Engineering, P.O. Box 577, Jhang Road, Faisalabad, Pakistan, 2004.

(10) Griffiths, K.; Partis, L.; Croan, D.; Wang, N.; Emslie, K. R. *Review of Technologies for Detecting Genetically Modified Materials in Commodities and Food*; Australian Government Analytical Laboratories, 2003.
(11) Blanco, M.; Villarroya, I. NIR spectroscopy: a rapid-response analytical tool. *Trends Anal. Chem.* **2002**, *21* (4), 240−250.
(12) Chen, J.; Arnold, M. A.; Small, G. W. Comparison of combination and first overtone spectral regions of near-infrared calibration models for glucose and other biomolecules in aqueous solutions. *Anal. Chem.* **2004**, *76* (18), 5405−5413.
(13) Sáiz-Abajo, M. J.; González-Sáiz, J. M.; Pizarro, C. Near infrared spectroscopy and pattern recognition methods applied to the classification of vinegar according to raw material and elaboration precess. *J. Near Infrared Spectrosc.* **2004**, *12*, 207−219.
(14) Chen, Q.; Zhao, J.; Zhang, H.; Liu, M.; Fang, M. Qualitative identification of tea by near infrared spectroscopy based on soft independent modeling of class analogy pattern recognition. *J. Near Infrared Spectrosc.* **2005**, *13*, 327−332.
(15) Cozzolino, D.; Smuth, H. E.; Gishen, M. Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. *J. Agric. Food Chem.* **2003**, *51*, 7703−7708.
(16) Andre, M. Multivariate Analysis and classification of the chemical quality of 7-Aminocephalsporanic acid using near-infrared reflectance spectroscopy. *Anal. Chem.* **2003**, *75*, 3460−3467.
(17) Gestal, M.; Gómez-Carracedo, M. P.; Andrade, J. M.; Dorado, J.; Fernández, E.; Prada, D.; Pazos, A. Classification of apple beverages using artificial neural networks with previous variable selection. *Anal. Chim. Acta* **2004**, *524*, 225−234.
(18) Casale, M.; Sáiz, Abajo, M. J.; González, Sáiz, J. M.; Pizarro, C.; Forina, M. Study of the aging and oxidation processes of vinegar samples from different origins during storage by near-infrared spectroscopy. *Anal. Chim. Acta* **2006**, *557*, 360−366.
(19) Roussel, S. A.; Hardy, C. L.; Hurburgh, C. R.; Rippke, G. R. Detection of roundup ready soybeans by near-infrared spectroscopy. *Appl. Spectrosc.* **2001**, *55*, 1425−1430.
(20) Rui, Y. K.; Luo, Y. B.; Huang, K. L.; Wang, W. M.; Zhang, L. D. Application of near-infrared diffuse reflectance spectroscopy to the detection and identification of transgenic corn. *Spectrosc. Spect. Anal.* **2005**, *25*, 1581−1583.
(21) Ying, T. J.; Yang, H. Q.; Xiang, Q. N.; Du, R. M. Characters of postharvest physiology of antisense *LeETR*1 transgenic tomato fruits. *Acta Hortic. Sin.* **2003**, *30* (4), 404−408.
(22) Hoel, B. O.; Solhaug, K. A. Effect of irradiance on chlorophyll estimation with the minolta SPAD-502 leaf chlorophyll meter. *Ann. Bot.* **1998**, *82*, 389−392.
(23) Lafargue, M. E.; Feinberg, M.; Daudin, J. J.; Rutledge, D. N. Detection of heterogeneous wheat samples using near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2003**, *11*, 109−121.
(24) Yan, Y. *Foundation of NIR Spectral Analysis and Its Application*; China Light Industry Press: Beijing, China, 2005.
(25) Geladi, P.; Macdougall, D.; Martens, H. Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Appl. Spectrosc.* **1985**, *10*, 491−500.
(26) Liu, Y.; Ying, Y.; Jiang, H. Rapid determination of maturity in apple using outlier detection and calibration model optimization. *Trans. ASABE* **2006**, *49* (1), 91−95.
(27) Wang, D.; Ram, M. S.; Dowell, F. E. Classification of damaged soybean seeds using near-infrared spectroscopy. *Trans. ASAE* **2002**, *45*(6), 1943−1948.
(28) Liu, Y. D.; Ying, Y. B.; Yu, H. Y.; Fu, X. P. Comparison of the HPLC method and FT-NIR analysis for quantification of glucose, fructose, and sucrose in intact apple fruits. *J. Agric. Food Chem.* **2006**, *54*, 2810−2815.
(29) Gray, J. E.; Picton, S.; Giovannoni, J. J.; Grierson, D. The use of transgenic and naturally-occurring mutants to understand and manipulate tomato fruit ripening. *Plant Cell Environ.* **1994**, *17* (55), 557−571.
(30) Cozzolino, D.; Vaz, Martins, D.; Murray, I. Visible and near infrared spectroscopy of beef longissimus dorsi muscle as a means of discriminating between pasture and corn feeding regimes. *J. Near Infrared Spectrosc.* **2002**, *10*, 187−193.

Received for review December 18, 2006. Revised manuscript received March 21, 2007. Accepted April 2, 2007. The authors gratefully acknowledge the financial support provided by the National Natural Science Foundation of China (Grants 30671197 and 30371001) and the Program for New Century Excellent Talents in University (No. NCET-04-0524).

JF063664M