Scientia Horticulturae 233 (2018) 99–107

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Quality assessment and discrimination of intact white and red grapes from *Vitis vinifera* L. at five ripening stages by visible and near-infrared spectroscopy

Hui Xiao[a], Ang Li[a], Meiyu Li[a], Ye Sun[a], Kang Tu[a], Shaojin Wang[b], Leiqing Pan[a,*]

[a] College of Food Science and Technology, Nanjing Agriculture University, Nanjing 210095, PR China
[b] Department of Biological Systems Engineering, Washington State University, 208 L.J. Smith Hall, Pullman, WA 99164-6120, USA

ARTICLE INFO

Keywords:
Total soluble solid contents
Total phenolic compounds
CIE L^* a^* b^*
Least square-support vector machine
Competitive adaptive reweighted sampling

ABSTRACT

The diffuse reflection visible/near-infrared (Vis/NIR, 400–1100 nm) and near-infrared (NIR, 900–2500 nm) spectrum were used to monitor the surface color (CIE $L^*a^*b^*$), total soluble solid contents (SSC) and total phenolic compounds (TP) of intact 'Manicure Finger' and 'Ugni Blanc' berries at five ripening stages (i.e., green, pre-veraison, veraison, post-veraison and ripe). The determination of quality parameters and the discrimination of five ripening stages were conducted by chemometric analysis based on full-band and selected wavelengths of Vis/NIR and NIR. The results showed that the best regression results were obtained by least squares support vector machine (LS-SVM) with the root mean squares error of prediction (RMSEP) of 5.161, 2.919, 3.275, 1.230% and 0.216 g kg^{-1} for L^*, a^*, b^*, SSC and TP of 'Manicure Finger' in the range of 400–1100 nm, respectively; and the RMSEP of 3.049, 0.710, 2.996 and 0.150 g kg^{-1} for L^*, a^*, b^* and TP of 'Ugni Blanc' in the range of 400–1100 nm, respectively, and the RMSEP of 1.288% for SSC in the range of 900–2500 nm. A total of 90% and 100% classification accuracies on prediction sets were reached by the total soluble solid contents based competitive adaptive reweighted sampling support vector machine discrimination analysis (SSC-based CARS SVM-DA) for 'Manicure Finger' and 'Ugni Blanc' grape berries of five ripening stages, respectively. This study provided a feasible evaluation method of quality and developing stages for grape varieties during ripening stages by Vis/NIR and NIR technology.

1. Introduction

Grape is a typical non-climacteric fruit where respiratory rates do not demonstrate a sharp increase after plucking from the grapevine. During the entire maturation process, the intensifying color of the berry peel is one of the most important and obvious changes for red grape (Owens, 2015), and the increase of the sugar content is generally appreciated to be the most important chemical changes in the grapes (Wang et al., 2010). The sugar content of the ripe fruit is closely related to the maturation stage and the rate of sugar accumulation (Huner et al., 1991). For red and black cultivars, the color of the skin is one of the most important parameters used to evaluate the quality and maturity of the berries, directly affecting the color and quality of the grape juice and wine (Ageorges et al., 2006; Yamamoto et al., 2015; Koyama et al., 2014).

The near infrared spectroscopy (NIR) technology, covering a spectral range of 780–2500 nm, combined with chemometric methods has found increasing applications in the detection of the ripening and the quality assessment of agricultural products. Organic molecules, such as sugar, organic acids and other chemical compounds, generally contain O—H, N—H, C—H and other hydrogen-bearing groups. The absorption bands of these structural components which related to overtone bands and combination of several stretch-bend vibration modes can be found in this spectral region (Qu et al., 2015).

In this context, one crucial goal is to identify the grapevine development stages for cultural and research uses in both fast and convenient ways. Non-destructive technologies were studied to address this goal by several researcher groups around the world. Fernández-Novales et al.

Abbreviations: Vis/NIR, visible/nearinfrared; NIR, near-infrared; SSC, total soluble solid contents; TP, total phenolic compounds; MS, moving-average smoothing; SNV, standard normal variate; PLS, the partial least squares; LS-SVM, the least-squares support vector machine; H_0, null hypothesis; SPA, successive projections algorithm; CARS, competitive adaptive reweighted sampling; R_c^2, determination coefficient of calibration; R_p^2, determination coefficient of prediction; RMSEC, root mean squares error of calibration; RMSEP, root mean squares error of prediction; RPD, the ratio of standard deviation to standard error; PCA, principle component analysis; PLS-DA, partial least squares discrimination analysis; SVM-DA, support vector machine discrimination analysis

* Corresponding author at: College of Food Science and Technology, Nanjing Agricultural University, NO.1 Weigang Road, Nanjing, 210095, PR China.
E-mail address: pan_leiqing@njau.edu.cn (L. Pan).

https://doi.org/10.1016/j.scienta.2018.01.041
Received 9 November 2017; Received in revised form 15 January 2018; Accepted 16 January 2018

(2009) studied the reducing sugar content during grape ripening by implementation of short wave and near infrared spectroscopy. Cirilli et al. (2015) developed a NIR-based method (1100–2300 nm) for the determination of physical and chemical changes during ripening stages of different olive cultivars and further explained firmness changes during the ripening process. Fourier-transform infrared spectroscopy was applied to study the qualitative evaluations of tartaric acid, malic acid, succinic acid, glucose and fructose, and quantitative discrimination of ripening stages for grapes in Musingarabwi et al.'s work (Musingarabwi et al., 2016).

Therefore, our work aimed to: (1) Comparison the performances of Vis/NIR and NIR spectra for the determination of grape quality (the surface color, total phenolic compounds and total soluble solid contents) during the ripening process; (2) Evaluate the feasibility of using spectral information combined with chemometrics to distinguish five ripening stages in comparison to the corresponding ripening parameter changes.

2. Materials and methods

2.1. Samples

In this study, *Vitis vinifera* L. cv. Manicure Finger and *Vitis vinifera* L. cv. Ugni Blanc were used for experiment. A total of 270 berry samples for each cultivar were collected from own-rooted 15-year-old vines (the vines were trained on overhead trellis with 2.5 m × 2.0 m spacing) in Nanjing Eight Diagrams Vineyard (32° 09′ 59.75″ N; 118° 49′ 29.82″ E, Nanjing, China) at five developmental stages (i.e., green, pre-veraison, veraison, post-veraison, ripe) during July (rainfall was 477.3 mm) and August (rainfall was 78.7 mm) of 2016 by experienced grower. Nine clusters from each stage of each cultivar and 6 berries were randomly picked from the top, middle and bottom of each cluster. The sampling process lasted for at least one month and the grape samples were collected between 8:00-10:00 am on each sampling day.

2.2. Visible/near-infrared and near-infrared system

After transported from the vineyard to the lab, all samples were allowed to equilibrate to room temperature (20 ± 0.5 °C) for 30 min before performing the experiment to eliminate any potential temperature effects. Then, the reflectance spectrum of each intact sample was measured by 32 scans using a fiber optic Vis/NIR system. The system consisted of a computer (Surface 3, Microsoft Corporation, USA), a visible-near-infrared fiber optic spectrometer (FX2000, Shanghai Ideaoptics Corporation, China) covering a spectral range between 400 and 1100 nm with 2048 individual wavelengths. A near-infrared fiber optic spectrometer (NIR2500, Shanghai Ideaoptics Corporation, China) operating at a spectral range between 900 and 2500 nm was used for measuring 256 wavelengths. Furthermore, a 7 strands-Y type optical fiber (FIB-Y-200-NIR, Shanghai Ideaoptics Corporation, China) with surrounding six strands for light emission and a central strand for acceptance was used. Further parts included a detection platform, a halogen light source (HL 2000, Shanghai Ideaoptics Corporation, Shanghai, China) with the bulb power for 9 W as well as a correction white board (STD-WS, Shanghai Ideaoptics Corporation, China) as reflective standard.

2.3. Chemical analyses

The reference parameters included the total soluble solid contents (SSC), surface color parameters in CIE $L^*a^*b^*$ space and total phenolic compounds (TP). The total soluble solid contents of each sample were measured by a digital hand-held 'pocket' refractometer (PAL-1, ATAGO, Japan), and were expressed in% with an accuracy of 0.1% unit. CIE $L^*a^*b^*$ describes one of the common color spaces and is based on the human perception of color. The factor L^* represents 'Lightness', with a value range from 0 to 100. The factor a^* represents the color change from magenta to green, and b^* represents the color change from yellow to blue. The CIE $L^*a^*b^*$ color space values were determined by a digital hand-held spectrophotometer (Ci6X, x-rite, USA).

The total phenolic compounds (TP) were determined to interpret the grape changes and were measured by a method previously used by Fragoso et al. (2011). The seeds of the samples were removed and the remaining tissue was homogenized in liquid nitrogen using a batch mill (A 11 B S025, IKA, Staufen Germany). Then, 0.2 g of the ground sample of each berry was macerated in 20 mL of hydroalcoholic acid solution (HCL 1% v/v: ethanol 96% = 17:3 v/v) at 40 °C for 40 min under ultrasonication. The mixture was then centrifuged for 5 min at 13,710 × *g* and the precipitate was resuspended in 20 mL of hydroalcoholic acid solution and centrifuged again. The supernatants were combined and diluted to 50 mL with hydroalcoholic acid solution. Finally, the absorbance of the extracted samples was measured at 280 nm using a UV spectrophotometer (UV 1800, Shimadzu Corporation, Japan) in order to determine the total phenolic compounds. Gallic acid monohydrate (in the range of 2.0–20.0 mg L^{-1}) was applied to generate a calibration curve and TP was expressed in gram (g) of gallic acid per kilogram (kg) of fresh grape sample.

All chemical analysis of each sample were conducted for three times, and the average were used as reference values for further analysis.

2.4. Determination of grape quality parameters

The spectral data of the samples were pre-processed by moving-average smoothing (MS) and the standard normal variate (SNV) to reduce noise and potential interferences of scatter and particle size (Chu et al., 2014).

After this process, the spectral data of each matrix were ranked by the reference values in descending order and every four samples constituted of one unit and the last sample of each unit was selected for prediction. All other three samples were selected for calibration. This sampling procedure resulted in 75% of the total samples from each variety for calibration and 25% for prediction (Pan et al., 2015).

The partial least squares (PLS) regression (Geladi and Kowalski, 1986) was used to establish regression models. A nonlinear regression method, the least-squares support vector machine (LS-SVM) (Suykens and Vandewalle, 1999) was also employed to predict the quality parameters in this study.

To verify the differences between two instruments which are statistically significant, Passing-Bablok regression (Passing and Bablok, 1983) was conducted on the prediction set, and based on the values of slopes and intercepts, the results of the best model were compared between the two NIR ranges and with the reference data (Malegori et al., 2017; Xiao et al., 2017). The slope and intercept were calculated with their 95% confidence interval. When the confidence interval for slope contains the value 1, then it is concluded that the hypothesis (slope = 1) is accepted; When the confidence interval for intercept contains the value 0, then the hypothesis (intercept = 0) is accepted. In conclusion, H_0 (null hypothesis) is accepted when slope = 1 and intercept = 0 are both accepted at a 95% confidence level.

Due to the enormous number of individual wavelengths obtained from the spectral data (the raw spectrum data of Vis/NIR contains 2048 wavelengths and the raw spectrum of NIR contains 256 wavelengths), a determination analysis was deemed to be too complex and time-consuming. Therefore, successive projections algorithm (SPA) (Araújo et al., 2001) and competitive adaptive reweighted sampling (CARS) (Li et al., 2009) were carried out to propose the most independent variables and develop a fast and accurate model. In this study, SSC-based and TP-based SPA and CARS were analyzed for the prediction of SSC and TP, respectively.

The performances of the regression models were evaluated by the determination coefficient of calibration (R_c^2), determination coefficient