Food Analytical Methods (2019) 12:936–946
https://doi.org/10.1007/s12161-018-01430-2

Near-Infrared Hyperspectral Imaging Rapidly Detects the Decay of Postharvest Strawberry Based on Water-Soluble Sugar Analysis

Qiang Liu[1] · Kangli Wei[1] · Hui Xiao[1] · Sicong Tu[2,3] · Ke Sun[4] · Ye Sun[5] · Leiqing Pan[1] · Kang Tu[1]

Received: 30 August 2018 / Accepted: 26 December 2018 / Published online: 10 January 2019

Abstract

This paper presents a novel strategy to detect the fungal decay in strawberry using reflectance near-infrared hyperspectral imaging (NIR-HSI, 1000–2500 nm). The variation of fructose, glucose, sucrose, and total water-soluble sugar (TWSS) content was analyzed using HPLC with a reference method during fungal infection in strawberry. The feasibility of quantifying sugar constituents relevant to the different stages of decay in strawberry was evaluated using NIR-HSI with key wavelengths selected via successive projection algorithm. The results showed that the predicted performance of TWSS content was acceptable within 0.807 for R_p^2 and 2.603 for RPD, respectively. Five to seven key wavelengths were obtained based on sugar constituents, and excellent performance for classification accuracy among the three stages of decay was 89.4 to 95.4% for calibration and 87.0 to 94.4% for prediction, respectively. This rapid approach provides a new strategy for the selection of key wavelengths to detect the decay and sugar constituents in strawberries.

Keywords Decay · Hyperspectral imaging · Key wavelength · Strawberry · Sugar content

Introduction

Strawberry (*Fragaria* × *ananassa Duch.*) is one of the most economically important fruits, popularly eaten fresh, processed into jams, and flavored for juices (Elmasry et al. 2007). Its content of the particular water-soluble sugars, including fructose, glucose, and sucrose, could directly affect the final taste and flavor of strawberry (Sturm et al. 2003). When strawberry ripens, it is typically harvested without stick peel as external protection, and easily contaminated with spoilage fungi during harvest and transportation. The presence of a few contaminated strawberry in a batch often causes cross-contamination, spreads to the entire batch of fruit, and eventually ruins the entire batch of strawberry (Cho et al. 2013). Therefore, an accurate and convenient system to monitor the health of strawberry to avoid substantial economic loss related to the microorganism contamination and reproduction should be developed.

Gray mold rot, caused by *Botrytis cinerea*, is one of the most serious diseases of strawberry. After the infection, the rot symptom is often accompanied by collapse and water soaking of parenchyma tissues, followed by a rapid appearance of gray masses of conidia on fruits. Then, sugars were gradually decomposed for the needs of fungi growth (Williamson et al. 2007). Several traditional direct methods have been developed to monitor fungi contamination, including microbiological cell counting and enzyme-linked immunosorbent assay (Thornton et al. 2010). However, the reagent preparation and processes of these methods are time-consuming and destructive to the fruit, and cannot be used in online detection. For these reasons, the development of fast, sensitive, and non-destructive approaches to detect the contaminated fungi in strawberry attracts commercial interest.

In recent years, a series of non-destructive methods have been developed to evaluate the chemical parameters of fruits. Near-infrared hyperspectral imaging (NIR-HSI) combines the

✉ Kang Tu
kangtu@njau.edu.cn

[1] College of Food Science and Technology, Nanjing Agricultural University, No. 1. Weigang Road, Nanjing, Jiangsu 210095, People's Republic of China

[2] Medical Sciences Division, University of Oxford, Oxford OX3 7BN, UK

[3] Sydney Medical School, The University of Sydney, Sydney, NSW 2006, Australia

[4] College of Environmental Science and Engineering, Anhui Normal University, No. 189. Jiuhua Road, Wuhu 241002, Anhui, China

[5] College of Engineering, Nanjing Agricultural University, No. 40. Dianjiangtai Road, Nanjing 210031, Jiangsu, China

nature of near-infrared spectroscopy with imaging into one system, enabling the provision of information on the distribution of major chemical constituent across individual samples (Wu and Sun 2013; Sun et al., 2017b). Indeed, some NIR-HSI calibrations have been successfully adopted into studies investigating the internal quality of thick rind fruits. Leiva-Valenzuela et al. (2014) used the hyperspectral imaging to evaluate the internal quality of blueberries. De Oliveira et al. (2014a) investigated the effectiveness of NIR-HSI to characterize the soluble solid contents (SSC) and titratable acidity (TA) in passion fruit; the good prediction performance with correlation coefficients of 0.93 and 0.95 for SSC and TA was obtained. Siedliska et al. (2018) reported the possibility of using NIR-HSI to detect the fungal infection in single strawberry. Although these studies have explored the feasibility of NIR-HSI system into the non-destructive evaluation of the fruit quality and characteristics, the vast amount data that generated from the studies takes time to calculate and process, limiting the speed and feasibility for online applications (Qin et al. 2013). From the earlier researches, Long et al. (2005) has successfully developed the expensive charge-coupled device (CCD) imaging system with NIR filters for the detection of total soluble sugars in melon tissue. However, the selection of wavelengths were constrained by the commercial availability (830, 850, 870, 905, and 930 nm), and the range for selection was restricted in 721 to 955 nm. Huang et al. (2017) has reported discrimination models on forecasting the days before decay of peach fruit based on selected variables. Hence, novel opportunities and extractions of featured data need to be investigated, which should accelerate the calculation speed to allow both spectral and spatial information processing to be implemented in a real-time system.

Sugar content is one of the most important indexes to evaluate the commercial values of strawberry since organoleptic assessment is greatly affected by the relative and total mounts of sugars (Darbellay et al. 2002). When sugar accumulation and conditions become favorable to disease development, dormant fungal infections could happen in strawberry later in the season at any time before or after harvest. Sturm et al. (2003) have confirmed that the sugar composition of this fruit varied remarkably among the stages of maturity; three sugar-related enzymes showed distinct activity patterns, as may affect sugar accumulation and the composition of strawberry (Basson et al. 2010). Several reports investigated and compared prediction of individual sugars (glucose, fructose, sucrose, and total sugars) in passion fruit and its pulp by using near-infrared and mid-infrared spectroscopic (De Oliveira et al. 2014a; Oliveira-Folador et al. 2018), and acceptable correlation coefficient values (R_v^2 with 0.942 for glucose, 0.855 for fructose and 0.818 for sucrose) were obtained. However, little is known about the connection between the sugar composition and fungal-caused decay. This linkage between sugar profiles and fungi diseases would be useful to develop a rapid analytical approach to assess the decay of fruits, as well as to evaluate the effectiveness for postharvest control treatment.

The strategy for monitoring fungal contamination in strawberry using reflectance NIR-HSI is presented here with the following goals: to (1) investigate the variation of sugar constituents during the fungal-infected decay of strawberry, (2) explore the feasibility of NIR-HSI for determination of sugar constituents in strawberry, (3) extract the featured wavelengths from the continued range of near-infrared region in the prediction of sugar constituents, and (4) provide a strategy to detect the different stages of decay based on sugar contents in strawberry.

Material and Methods

Sample Preparation

The research was carried out using one of the most popular strawberry (*Fragaria* × *ananassa* Duch.) variates in south China, namely '*Hongyan*'. Samples were purchased from the local strawberry orchard in Nanjing (China, 32° 07′ 35″ N, 118° 59′ 27″ E) on 1 February 2018. Strawberries, 45 days after flowering, were harvested at commercial maturity and were visually inspected to be absent of surface bruise or fungal contamination. After picking from the orchard, all strawberries were immediately transported to the lab. The samples were immersed in 0.1% (*V*/*V*) trichloroacetic acid for 2 min and then rinsed twice with sterile distilled water.

In the fungal infection experiment group, strawberries were inoculated with 20 μL of *Botrytis cinerea* spore solution (1×10^5 spores mL^{-1}, its concentration was measured using a hemocytometer, and adjusted to final concentration with sterile saline solution) at a 2-mm depth, in contrast to the sterile saline solution as the control group (the CK group). When moisture was evaporated at room temperature (20 ± 5 °C), the strawberries were stored at 20 °C under 85% relative humidity. During the storage, samples were taken at various time points to detect the changes of decay. An external experiment was taken into consideration to verify the accuracy of the model for rapid detection of infected strawberries. At each time point, (1) one group including 15 strawberries was adopted to obtain their hyperspectral imaging datasets, and the data were used to analyze their water-soluble sugar content, and (2) another group including 54 samples were randomly selected from the datasets of hyperspectral imaging and growth character of the fungi on sample surface for external validation. Each sample was directly measured (NIR-HSI) and then was frozen by liquid nitrogen and stored at − 24 °C for the further biochemical analyses.

Sugar Analysis

Sugar profile analysis and quantification were performed by an HPLC system (Shimadzu LC-20A, Kyoto, Japan) according to the procedures described by Yu et al. (2016) with some modifications.

In the measurement of sugar compounds, the slices with a nominal thickness (1 cm, measured by digital vernier caliper) from the scanned surface in sample were cut and then ground into powder in liquid nitrogen. For water-soluble sugar extraction, 5.000 ± 0.100 g of sample powder was weighted and was added with 30 mL of ultrapure water and equilibrated in a water bath at 80 °C for 1 h. The tube was centrifuged at 10,000 rpm for 20 min at 20 °C. A volume of 1 mL supernatant solution was distilled with 4 mL water, and filtered through a 0.45-μm membrane filter. Each measurement was conducted in triplicate.

The HPLC procedures were set as the following: 4.6 × 250 mm zorbaxcarbohydrate column (Agilent Technologies Inc., CA, USA), the mobile phase of 75% acetonitrile in ultrapure water at flux 1.0 mL min^{-1}, column temperature at 40 °C, and running time at 18 min. The injected sample volume was 10 μL. The evaporative light-scattering detector (ELSD) was used for analysis of sugar compounds where UV detection might be a restriction, and therefore, compounds do not efficiently absorb UV radiation (Douville et al. 2006). Computer-assisted optimization program was used for data analysis. The external standard method was adopted by using D(−)-fructose, D(+)-glucose, and sucrose at analytical grade with > 98% purity (Yuanye Bio-Technology Co., Ltd., Shanghai, China) as standards, respectively. Three calibration curves (R^2 = 0.9995, RMSE = 0.007 mg g^{-1} for fructose; R^2 = 0.9993, RMSE = 0.024 mg g^{-1} for glucose; R^2 = 0.9998, RMSE = 0.004 mg g^{-1} for sucrose) were established for fructose, glucose, and sucrose, respectively. The sum of fructose, glucose, and sucrose was expressed as total water-soluble sugar (TWSS). The unit was expressed as milligram TWSS over gram of fresh fruit weight.

NIR-HSI System

The setup of NIR-HSI system mainly consists of a light source, a near-infrared spectral camera (Xeva-T2SL-9928, Xenics Ltd., Leuven, Belgium), a transport unit, and several optical fibers. Each lamp controller consists of three 50-W halogen lamps and was fixed in the opposite frames positioned at the angle of 60° towards the transport unit surface. The spectral camera includes a cooled 320 × 256 pixel (14 bits) mercury cadmium telluride (HgCdTe) detector and an N25E imaging spectrometer. The lens is with a focal length of 31 mm and focal distance ratio of 2.0 (number 031505, Specim Ltd., Oulu, Finland). The transport unit consists of a positioning unit and an electronic-motorized horizontal stage (IRCP0076, Isuzu, Zhubei, China).

The line scanned, and reflectance mode, was adopted in this experiment. The distance from light lamps to samples was 30 cm. The speed of the transport unit was set at 13.78 mm s^{-1}. The exposure time was adjusted to 2.5 ms. The spectral resolution was approximately 6.8 nm from 1000 to 2500 nm (234-dimension vectors). The imaging data was saved as the RAW format in the computer (No. Vostro-3667 (i5-6400, 8G memory), Dell, TX, USA) for subsequent analysis.

Imaging Processing and Data Extraction

Automatic calibration was conducted using the dark and white reference images in both spectral and spatial dimensions using the following equation (Caporaso et al. 2018):

$$I_{cal} = \frac{I_i - I_d}{I_w - I_{id}} \frac{t_w}{t_s}$$

where I_{cal} is the calibrated relative reflectance of raw images, I_i is the raw images, I_d is the dark image value for white calibration (reflectance ~ 0%), I_w is the whit reference value (reflectance ~ 99%, PTFE material), I_{id} is the dark image value for sample calibration, t_w is the sample exposure time, and t_s is the white exposure time. The calibration process was measured of the white reference with 50 repeats. To extract the mean spectral of the strawberry samples, an image mask procedure was applied. The mean spectra for the region of the individual strawberry were automatically calculated and exported for statistical analysis.

Data Processing

Methodology applied for data analysis was accordant with Oliveira et al. (2014a). Spectral pre-processing methods were used to remove the effects of unflatten surface of the samples. Standard normal variate correction (SNV) was used to compensate for baseline shifting, non-uniform scattering, and other physical effects. This methodology was also recommended for the practical application (Nicolaï et al. 2007; Miguel and Flor 2013; Felix et al. 2018).

Prediction Model Analysis

A support vector machine (SVM) regression was built using the mean spectra obtained from the individual strawberry sample and the corresponding chemical measurements of fructose, glucose, sucrose, and TWSS using HPLC. In this work, calibration models with leave-one-out cross-validation were used in calibration step. Leave-one-out cross-validation is useful because it does not waste data and it is more suited to the

small amount of samples (De Oliveira et al. 2014b; Liu et al. 2018). In order to verify the performance of prediction models for quality attributes, a population of 60 samples (harvested in 12 March 2018) was defined as external validation. These external validation samples were not used in calibration and cross-validation steps, in order to avoid overfitting as suggested by De Oliveira et al. 2014a

The best SVM model was chosen according to the correlation coefficient of calibration (R_c^2), the correlation coefficient of cross-validation (R_{cv}^2), the correlation coefficient of prediction (R_p^2), the root mean square error of calibration (RMSEC), the root mean square error of prediction (RMSEP), the root mean square error of cross-validation (RMSECV), and the relative percentage of deviation (RPD), respectively. The optimal parameters of SVM were chosen based on the RMSECV to minimize the prediction error and reduce the potential risk of overfitting. The parameters were defined as the following:

$$R_C^2, R_{CV}^2, R_P^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \overline{y}_i\right)^2}{\sum_{i=1}^{n} \left(y_i - y_m\right)^2}$$

$$\mathrm{REMSEC}, \mathrm{REMSECV}, \mathrm{RMSEP} = \sqrt{\frac{1}{n} \sum_{I=1}^{n} \left(y_i - \overline{y}_i\right)^2}$$

$$\mathrm{RPD} = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left(y_i - y_m\right)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \overline{y}_i\right)^2}}$$

where n is the total number of samples, y_i and $\overline{y}$ are the measured and the predicted values of the ith sample, and y_m is the mean value of all samples in calibration, cross-validation, or prediction set, respectively. The statistic parameter RPD is desired to be larger than 2 for a good prediction.

Selection of Key Wavelengths with Successive Projection Algorithm

The successive projection algorithm (SPA) is a forward selection method that applies projection operations in a vector space to obtain subsets of variables with small collinearity (Araújo et al. 2001). Regression models were built for the prediction of fructose, glucose, sucrose, and TWSS, using the spectral data as matrix X and reference data by HPLC as matrix Y for the calibration. The description and main procedures of SPA have been detailed in a previous report (Araújo et al. 2001). In our report, the input layer contained a total of 90 reflected spectra of strawberry samples and their corresponding fructose, glucose, sucrose, and TWSS levels. The maximum number of variables (potential key wavelengths) was selected using SPA as 15 based on our pre-experiment results.

Visualization Image of Water-Soluble Sugar Level

Visualization of the detailed spectral and spatial information from its hypercube reveals the discrepancies in the quality attributes from one sample to another sample (Caporaso et al. 2018). After the extraction of wavelengths from the raw NIR-HSI data, they were used to make perditions with less time-consuming and lower calibration cost. The contents of water-soluble sugar in each pixel of the spectral image on strawberry samples were assessed using the prediction model. The assessing images were then used to obtain the pseudo-colors with the scales shown as "visualization images."

Classification of Fungal-Caused Decay in Strawberries

The stage of fungal decay was visually inspected and measured according to a previous report (Pan et al. 2014) with some modifications. Based on the growth character of the *B. cinerea* mold on the sample surface, three infected stages were given to individual strawberry, where in the first stage, less than 5% of the sample surface were infected; in the second stage, 5–20% of the sample surface were infected; and in the third stage, more than 20% of the sample surface were infected. The results were used to evaluate the fungal growth period on strawberries. Then, classifying the stages of decay was performed based on the selected hyperspectral data of water-soluble sugar contents in strawberries.

Statistical Analysis

Statistical analysis was performed using the SPSS 18.0 software (IBM Corporation, Armonk, NY, USA). The least significant difference (LSD) between quality data of each sample group during storage period was evaluated using one-way analysis of variance (ANOVA). The difference at a level of 0.05 ($P < 0.05$) was considered significant. The graphs were prepared using the Origin 9.0 software (Origin Lab Corporation, Northampton, MA, USA).

The processing and construction of the SVM models were carried out using the PLS Toolbox 7.5 (Eigenvector Research, Inc., Wenatchee, WA, USA), which was run on MATLAB 2014 (The MathWorks Inc., Natick, MA, USA). A program written in MATLAB was applied to perform SPA selection. Each experiment was repeated three times, and the mean values were used for data processing.

Results and Discussion

Analysis of Water-Soluble Sugar Content

The contents of fructose, glucose, sucrose, and TWSS in strawberries determined using the HPLC method are

presented in Fig. 1. The changes in sugar profiles were observed according to the time points of storage and fungal infection. Sucrose was the most abundant soluble sugar in healthy strawberries at 0 h. The maximum contents were two and five times higher than those of fructose and glucose, respectively. After the fungal infection of strawberries, the differences in sugar contents significantly decreased, and the content of sucrose was less than that of fructose at 96 h of storage (Fig. 1a). In the CK group, the mean values of sugar contents were little changed, being less than 10 mg g^{-1}. The sucrose remained as the primary soluble sugar in samples (Fig. 1b). During the storage from 0 to 120 h at 20 °C, the average TWSS content decreased from 69.33 ± 4.64 to 61.41 ± 6.47 mg g^{-1} in the fungal-infected groups. The decline in TWSS content was mainly due to the significant reduction of sucrose content (from 38.76 ± 8.05 to 4.50 ± 0.98 mg g^{-1}; $P < 0.01$), which contributes to 81.07% of TWSS reduction. In contrast, TWSS in the CK group was relatively stable, ranging from 69.33 ± 4.64 to 61.41 ± 6.47 mg g^{-1}. The levels of sugar component ranged in the following order: sucrose > fructose > glucose during the 120-h storage period.

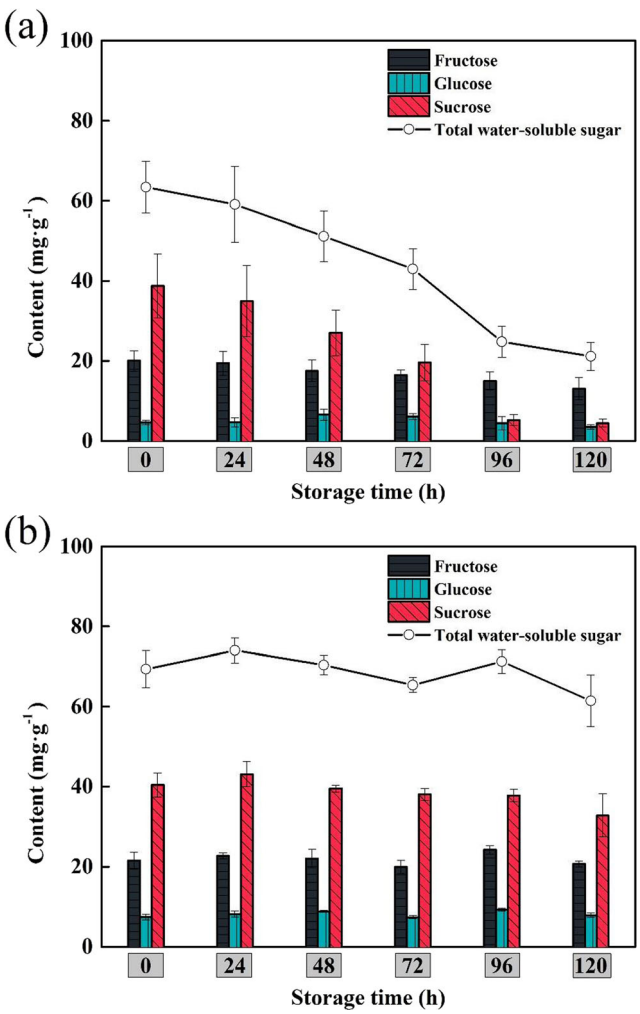

Fig. 1 HPLC for the analysis of fructose, glucose, sucrose, and total water-soluble content in strawberries during different storage periods. **a** Fungal-infected group; **b** control group. TWSS represents total water-soluble sugar content. Data were express as mean ± standard error

Our results of the natural variability of fructose, glucose, sucrose, and TWSS in '*Hongyan*' strawberry are in agreement with their average values reported previously (Chen 2007), however lower than the glucose levels in another study (Sturm et al. 2003) that reported a broad range from 16.3 to 28.2 mg g^{-1} in 13 cultivars of ripen strawberries. The variation of the sugar contents in strawberries is justifiable since it could be affected by several factors, including species, location, maturity, and the picking period. Zhang (2017) has reported that the chemical composition of CVS. '*Hongyan*' depends on the maturity stage, and sugar accumulation is dominated by the increasing of sucrose content which could reach as high as 50.74 mg g^{-1} in CVS. '*Hongyan*'. Meanwhile, the contents of total water-soluble sugar (~ 72 mg g^{-1} FW) and sucrose (~ 43 mg g^{-1} FW) in fruits harvested in February 2004 are at the highest level, followed by January, whereas the lowest total water-soluble sugar level (~ 40 mg g^{-1} FW) has been observed in April (Chen 2007). These observations are in agreement with our findings, suggesting that the sugar composition of strawberries could be varied among the genotypes of the plant and other factors. Nevertheless, the sugar composition of the same batch of healthy strawberries should be stable under the same controlled storage condition.

We found that the content of TWSS deceased by 12.28 mg g^{-1} during the first 48 h of storage in the fungal-infected group. TWSS level further declined by 26.37 mg g^{-1} from 48 to 96-h time point. Correlation coefficients between the stages of decay and sugar compounds are shown in Table 1. The stages of decay were significantly correlated with the sucrose and TWSS

Table 1 Correlation coefficients between sugar compounds and stages of fungal decay

	Decay stage	Fructose	Glucose	Sucrose	TWSS
Decay stage	1	−0.668*	−0.193	−0.899**	−0.909**
Fructose	–	1	0.418	0.575*	0.720**
Glucose	–	–	1	0.231	0.361
Sucrose	–	–	–	1	0.978**
TWSS	–	–	–	–	1

Correlation values followed by * are significant at $P < 0.05$ and ** are significant at $P < 0.01$

TWSS total water-soluble sugar, the sum of fructose, glucose, and sucrose

content ($P < 0.01$), respectively. In addition to the glucose level, contents of fructose, sucrose, and TWSS in the individual sample also showed significant correlations ($P < 0.05$). These results suggest that the water-soluble sugar consumption might offer the opportunity to extract the characteristics of fruit decay progression and make monitoring the fungal infection process in strawberry possible.

Spectral Analysis

The values of relative reflectance to individual strawberry, including infected and control groups, were calculated and shown in Fig. 2a, c. The average values of the mean spectra from 0 to 120 h are shown in Fig. 2b, d. The main absorption peaks were at 1160, 1410, 1760, 1900, and 2300 nm, respectively. Those absorption wavelengths were related to the vibration and combination overtones of the *C–H*, *O–H*, and *N–H* chemical bonds which are the primary structural components of organic molecules (Cevoli et al. 2013). The peak at 1160 nm was due to the strong absorption of carbohydrates (Siedliska et al. 2018). The water absorption peaks were observed at 1400 and 1900 nm, respectively, which were related to the *O–H* fist overtone and *O–H* stretch deformation, respectively (Williams 1987). The absorption peak at 1760 nm was related to the liquid that was attributed to the $-CH_2$ overtone (Huck et al. 2005). Another minor absorption peak was observed at 2300 nm which was associated with *C–H* stretching (Ma et al. 2018).

Figure 2b demonstrates the lower absorption values in the fungal-infected group compared with the normal strawberries during storage. The reason can be explained as the release of the water into the apoplast of diseased samples, and then, the air spaces filled with water-causing tissue become more translucent, and lower light absorption was observed. A previous study has demonstrated that the most prominent absorption bands in the NIR region were related to *C–H* and *O–H* functional bonds, and these bonds are common in major constituents in water, sucrose, and cellulose within the samples (Cevoli et al. 2013). In our experiments, precise values of the water-soluble sugar compounds of the individual sample were determined by HPLC. Further analysis using multivariate datasets was carried out to establish

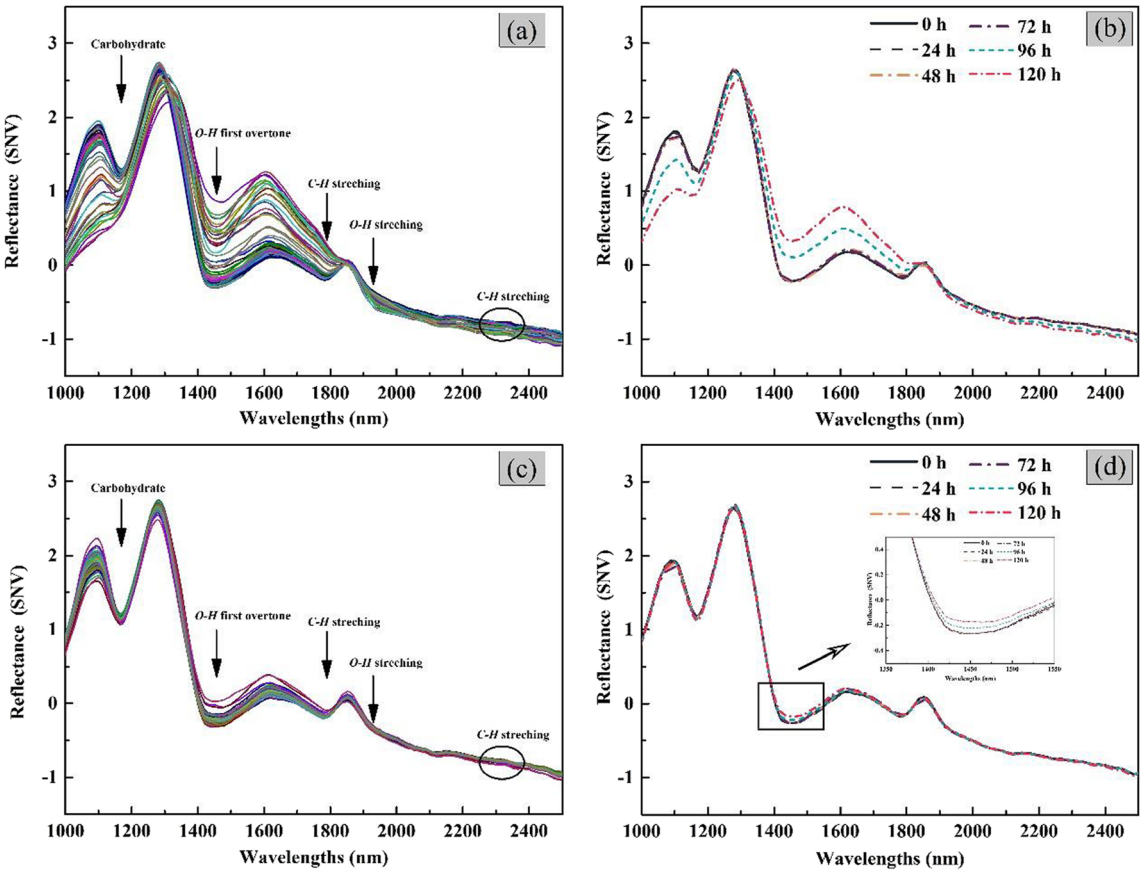

Fig. 2 The relative reflectance spectra of strawberries obtained using near-infrared hyperspectral imaging system from 1000 to 2500 nm. **a** Individual strawberry; **b** mean spectra during each storage period

Table 2 Performance of the regression models for fructose, glucose, sucrose, and total water-soluble sugar content of individual strawberry

Dataset	SVM parameters		R_c^2	RMSEC	R_{cv}^2	RMSECV	R_p^2	RMSEP	RPD
	Gamma	Cost		(mg g^{-1})		(mg g^{-1})		(mg g^{-1})	
Fructose	0.00032	100	0.702	1.796	0.608	2.049	0.589	2.307	1.532
Glucose	0.032	1	0.551	1.097	0.504	1.143	0.502	1.180	1.371
Sucrose	0.0032	31.62	0.774	4.386	0.741	5.630	0.724	5.190	2.185
TWSS	0.0032	10	0.824	5.530	0.803	5.532	0.807	5.561	2.603

R_c^2, R_{cv}^2, and R_p^2: the correlation coefficient of calibration, cross-validation, and prediction, respectively

SVM support vector machine, RMSEC root mean square error of calibration, *RMSECV* root mean square error of cross-validation, *RMSEP* root mean square error of prediction, *TWSS* total water-soluble sugar, the sum of fructose, glucose, and sucrose, *RPD* residual predictive deviation

;

the quantitative prediction models, and the key wavelengths were selected based on the sugar content.

Quantification and Visualization of Water-Soluble Sugar Content

Table 2 reports the performance of the SVM regression models for the prediction of fructose, glucose, sucrose, and TWSS in individual strawberry using the NIR-HSI datasets. As described in above ("Prediction Model Analysis" section), the values of R^2, RMSE, and RPD were calculated to evaluate the performance of the prediction models. In the modeling process, the leave-one-out cross-validation was used to get the best model parameters (gamma and cost values) in SVM models. Compared with three individual sugar compounds, namely fructose, glucose, and sucrose, the sucrose model performed considerably better RMSECV values (5.630 mg g^{-1}), with 0.741 R_{cv}^2. The model built on the TWSS was deemed to be the best prediction model because it used all the features from the contents of these three sugars with a higher R_{cv}^2 (0.803).

Generally, the RPD values higher than 2 could be considered as the good prediction. For sucrose and TWSS, the RPD characterized high value for prediction ability: RPD for TWSS was 2.603. The predicting performance showed that the model offered high accuracy for TWSS prediction based on NIR-HSI dataset. Therefore, the "chemical distribution images" were generated and visualized to reveal the distribution of TWSS compounds in a single sample. Figure 3 shows the TWSS distribution on the tested samples from 0 to 120-h storage time points. Once the prediction and visualization were applied, the discrepancies of the TWSS were visible between various storage periods. Strawberries had significantly higher TWSS contents with a more uniform distribution during 0–48 h than TWSS contents during the 72–120 h. When the visualization was adopted to predict the TWSS content, a slight difference between 0 and 48-h period was visualized. Because NIR-HSI can provide both the content and distribution information for the individual strawberry, a potential online system for rapid scanning could be developed based on this visualized approach.

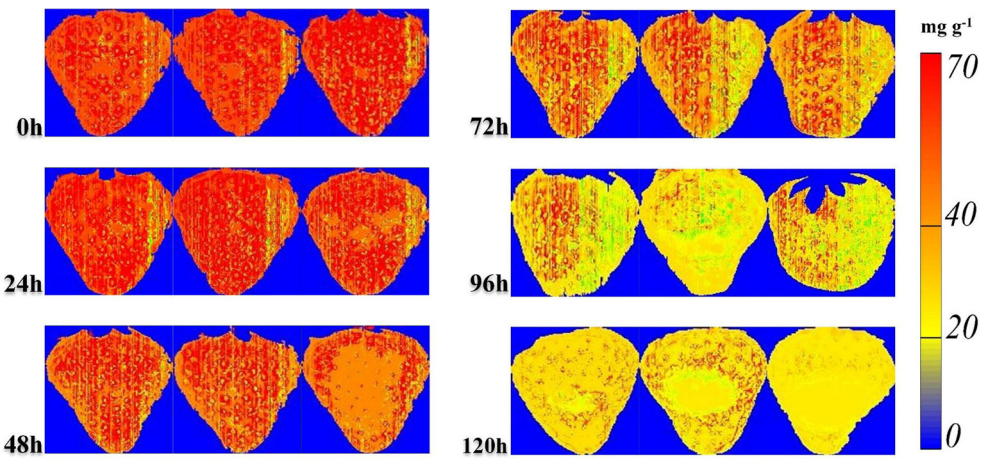

Fig. 3 Application of the calibrations to visualize total water-soluble sugar content in individual strawberry

Table 3 Key wavelength selection based on fructose, glucose, sucrose, and total water-soluble sugar content datasets

	Number*	Key wavelengths	Common wavelengths
Fructose	5	1085 nm, 1158 nm, 1348 nm, 1428 nm, 1915 nm	1158 nm; 1348 nm; 1428 nm
Glucose	5	1158 nm, 1193 nm, 1348 nm, 1428 nm, 1516 nm	
Sucrose	5	1158 nm, 1234 nm, 1348 nm, 1428 nm, 1590 nm	
TWSS	7	1085 nm, 1158 nm, 1234 nm, 1348 nm, 1486 nm, 1590 nm, 1780 nm	

TWSS total water-soluble sugar, the sum of fructose, glucose, and sucrose

*Number of key wavelengths

Key Wavelength Selection with Successive Projection Algorithm

A total of 90 samples and their related sugar compounds, namely fructose, glucose, sucrose, and TWSS, were generated using the selected models using key wavelengths. The more key wavelengths were selected for modeling, the more calculation time was required, and slower calculation speed was achieved. Based on the RMSE results of key variable selection after SPA processing, the values remained mostly stable after ranking at 5 for fructose, glucose, and sucrose and 7 for TWSS (Table 3). As described above in "Spectral Analysis" section, these wavelengths are at the absorption peak of carbohydrates related to *C–H* or *O–H* chemical bonds (Siedliska et al. 2018). Besides, the special bands such as 1780 nm were selected based on TWSS, which is related to the $-CH_2$ overtone (Huck et al. 2005). This effect was because the TWSS content was determined by the total of fructose, glucose, and sucrose contents in strawberry. The results of predicting performance based on selected wavelengths by SVM are shown in Table 4. The SPA-SVM models based on fewer variables can achieve good accuracy to predict TWSS content. In comparison with the results of the full-range NIR-HSI spectra, the SPA-SVM model revealed relatively low accuracy for prediction. The reason is that those key wavelengths were selected for the primary characteristics rather than the whole raw data. Similar results have been reported by previous studies predicting chlorophyll content in peaches (Sun et al. 2017a).

Classifying Three Stages of Decay in Strawberry During Storage by Key Wavelengths

The external dataset with additional 324 strawberries (three stages of decay using 108 samples at each stage) was obtained using the NIR-HSI. The classification experiments for stages of fungal decay were performed based on the key wavelengths, aiming to verify the possibility of further rapid and online detection of contaminated strawberries. The final results of the four classification models for distinguishing fungal decay at the first stage, second stage, and third stage are shown in Table 5. The best prediction for evaluating the fungal decay was obtained from the key wavelengths based on TWSS (95.4% accuracy for calibration sets; 94.4% for prediction sets). The other models obtained considerably lower accuracy being 89.4 to 94.0% for calibration sets and 87.0 to 92.6% for prediction sets. As the calibration results shown in Table 5, identifying the second stage of decay could be possibly confused with other infection stages. These demonstrations showed that the NIR-HSI technology could be successfully applied to identify the stages of decay in contaminated strawberry during storage. Furthermore, the results indicated that the SPA methods combined with water-soluble sugar composition analysis offered the substantial potential to monitor the stages of decay in strawberries. These methods were also useful for the further design of an online system to identify other types of decay fruits.

Table 4 Performance of the regression models for fructose, glucose, sucrose, and total water-soluble sugar content based on selected wavelengths

Dataset	SVM parameters		R_c^2	RMSEC	R_{cv}^2	RMSECV	R_p^2	RMSEP	RPD
	Gamma	Cost							
Fructose	0.32	10	0.688	1.830	0.584	2.115	0.572	2.332	1.516
Glucose	0.32	0.32	0.531	1.318	0.469	1.376	0.498	1.262	1.282
Sucrose	1	31.62	0.770	4.626	0.746	5.522	0.720	5.755	1.971
TWSS	1	3.16	0.810	5.678	0.782	6.379	0.774	6.459	2.244

R_c^2, R_{cv}^2, and R_p^2: the correlation coefficient of calibration, cross-validation, and prediction, respectively

SVM support vector machine, *RMSEC* root mean square error of calibration, *RMSECV* root mean square error of cross-validation, *RMSEP* root mean square error of prediction, *TWSS* total water-soluble sugar, the sum of fructose, glucose, and sucrose, *RPD* residual predictive deviation

Table 5 Identification of fungal-infected phase in individual strawberry using SVM model

Key wavelengths	Calibration				Prediction			
		I	II	III		I	II	III
Fructose	I	70	4	0	I	36	6	0
	II	2	64	4	II	0	29	1
	III	0	4	68	III	0	1	35
	Accuracy	97.2%	88.9%	94.4%	Accuracy	100.0%	80.6%	97.2%
	Total		93.5%		Total		92.6%	
Glucose	I	66	4	0	I	30	4	0
	II	6	61	6	II	6	31	3
	III	0	7	66	III	0	1	33
	Accuracy	91.7%	84.70%	91.7%	Accuracy	83.3%	86.1%	91.7%
	Total		89.4%		Total		87.0%	
Sucrose	I	72	4	0	I	30	3	0
	II	0	65	6	II	6	32	3
	III	0	3	66	III	0	1	33
	Accuracy	100.0%	90.3%	91.7%	Accuracy	83.3%	86.1%	91.7%
	Total		94.00%		Total		87.0%	
TWSS	I	72	4	0	I	34	4	0
	II	0	65	3	II	2	32	0
	III	0	3	69	III	0	0	36
	Accuracy	100.0%	90.3%	95.8%	Accuracy	94.4%	88.9%	100.0%
	Total		95.4%		Total		94.4%	

I, II, and III: first, second, and third stage of fungal decay, respectively

SVM support vector machine, *TWSS* total water-soluble sugar, the sum of fructose, glucose, and sucrose

Conclusion

This paper introduces the possibility of using NIR-HSI to non-destructively quantify the chemical sugar constituents of strawberry, and the potential to monitor the stages of fungal decay during storage. Statistical results showed that the changes at the individual sugar level of strawberries were mainly affected by the infection period from 0 to 120 h. This NIR-HSI method offers the rapid prediction of sugar components, especially the TWSS content with high RMSECV and RPD for TWSS being 5.561 mg g^{-1} and 2.603, respectively.

The NIR-HSI is a suitable method to predict the variability of sugar constituents in strawberry. Key wavelengths in near-infrared region were selected by SPA based on sugar content to eliminate the redundant and interpretation information obtained from NIR-HSI analysis. The outcomes of R_p^2 range from 0.498 to 0.720 for individual sugar quantification. Moreover, the new strategy using key wavelengths based on sugar content was developed to monitor the stages of decay in strawberries during storage. The performance of classification accuracy among the decay stages was 95.4% for calibration and 94.4% for prediction based on the TWSS-selected wavelengths, respectively. These findings suggest that the NIR-HSI is a new promising technology to detect and monitor possible fungal decay and sugar constituents during strawberry storage.

Funding Information This study is financially supported by National Natural Science Foundation of China (NSFC 31671926; 31671925), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and 2017 Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX17_0631).

Compliance with Ethical Standards

Conflict of Interest Qiang Liu declares that he/she has no conflict of interest. Kangli Wei declares that he/she has no conflict of interest. Hui Xiao declares that he/she has no conflict of interest. Sicong Tu declares that he/she has no conflict of interest. Ke Sun declares that he/she has no conflict of interest. Ye Sun declares that he/she has no conflict of interest. Leiqing Pan declares that he/she has no conflict of interest. Kang Tu declares that he/she has no conflict of interest.

Ethical Approval This article does not contain and studies with human participants or animals performed by any of the authors.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Araújo MCU, Saldanha TCB, Galvão RKH, Yoneyama T, Chame HC, Visani V (2001) The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. Chemom Intell Lab Syst 57(2):65–73. https://doi.org/10.1016/S0169-7439(01)00119-8

Basson CE, Groenewald JH, Kossmann J, Cronjé C, Bauer R (2010) Sugar and acid-related quality attributes and enzyme activities in strawberry fruits: invertase is the main sucrose hydrolysing enzyme. Food Chem 121(4):1156–1162. https://doi.org/10.1016/j.foodchem.2010.01.064

Caporaso N, Whitworth MB, Grebby S, Fisk ID (2018) Non-destructive analysis of sucrose, caffeine and trigonelline on single green coffee beans by hyperspectral imaging. Food Res Int 106:193–203. https://doi.org/10.1016/j.foodres.2017.12.031

Cevoli C, Gori A, Nocetti M, Cuibus L, Caboni MF, Fabbri A (2013) Ft-nir and ft-mir spectroscopy to discriminate competitors, non-compliance and compliance grated parmigiano reggiano cheese. Food Res Int 52(1):214–220. https://doi.org/10.1016/j.foodres.2013.03.016

Chen JW (2007) Difference in sugar content of fruit harvested in different month strawberry (fragaria×ananassa duch'tochiotome') and its relation to sucrose metabolism. Acta Horticulturae Sinica 34(5):1147–1150. https://doi.org/10.16420/j.issn.0513-353x.2007.05.017

Cho BK, Kim MS, Baek IS, Kim DY, Lee WH, Kim J, Bae H, Kim YS (2013) Detection of cuticle defects on cherry tomatoes using hyperspectral fluorescence imagery. Postharvest Biol Technol 76:40–49. https://doi.org/10.1016/j.postharvbio.2012.09.002

Darbellay C, Carlen C, Azodanlou R, Villettaz JC (2002) Measurement of the organoleptic quality of strawberries. Acta Hortic 567:819–822. https://doi.org/10.17660/ActaHortic.2006.701.100.

De Oliveira GA, Bureau S, Renard CMGC, Pereira-Netto AB, de Castilhos F (2014a) Comparison of NIRS approach for prediction of internal quality traits in three fruit species. Food Chem 143:223–230. https://doi.org/10.1016/j.foodchem.2013.07.122

De Oliveira GA, de Castilhos F, Renard CMGC, Bureau S (2014b) Comparison of NIR and MIR spectroscopic methods for determination of individual sugars, organic acids and carotenoids in passion fruit. Food Res Int 60:154–162. https://doi.org/10.1016/j.foodres.2013.10.051

Douville V, Lodi A, Miller J, Nicolas A, Clarot I, Prilleux B et al (2006) Evaporative light scattering detection (elsd): a tool for improved quality control of drug substances. Pharmeur Sci Notes 2006(1):9–15

Elmasry G, Wang N, Elsayed A, Ngadi M (2007) Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. J Food Eng 81(1):98–107. https://doi.org/10.1016/j.jfoodeng.2006.10.016

Felix YHK, Chen Q, Hassan MM, Yang M, Sun H, Rahman MH (2018) Near infrared system coupled chemometric algorithms for enumeration of total fungi count in cocoa beans neat solution. Food Chem 240:231–238. https://doi.org/10.1016/j.foodchem.2017.07.117

Huang L, Meng L, Zhu N, Wu D (2017) A primary study on forecasting the days before decay of peach fruit using near-infrared spectroscopy and electronic nose techniques. Postharvest Biol Technol 133:104–112. https://doi.org/10.1016/j.postharvbio.2017.07.014

Huck CW, Guggenbichler W, Bonn GK (2005) Analysis of caffeine, theobromine and theophylline in coffee by near infrared spectroscopy (nirs) compared to high-performance liquid chromatography (hplc) coupled to mass spectrometry. Anal Chim Acta 538(1):195–203. https://doi.org/10.1016/j.aca.2005.01.064

Leiva-Valenzuela GA, Lu R, Aguilera JM (2014) Assessment of internal quality of blueberries using hyperspectral transmittance and reflectance images with whole spectra or selected wavelengths. Innov Food Sci Emerg Technol 24:2–13. https://doi.org/10.1016/j.ifset.2014.02.006

Liu Q, Zhao N, Zhou D, Sun Y, Sun K, Pan L, Tu K (2018) Discrimination and growth tracking of fungi contamination in peaches using electronic nose. Food Chem 262:226–234. https://doi.org/10.1016/j.foodchem.2018.04.100

Long RL, Walsh KB, Greensill CV (2005) Sugar "imaging" of fruit using a low cost charge-coupled device camera. J Near Infrared Spectrosc 13(4):177–186. https://doi.org/10.1255/jnirs.536

Ma T, Li X, Inagaki T, Yang H, Tsuchikawa S (2018) Noncontact evaluation of soluble solids content in apples by near-infrared hyperspectral imaging. J Food Eng 224:53–61. https://doi.org/10.1016/j.jfoodeng.2017.12.028

Matlab U. S. G. (2014) The mathworks. Inc., Natick, MA, 2014

Miguel M, Flor AT (2013) Spectroscopic determination of aboveground biomass in grasslands using spectral transformations, support vector machine and partial least squares regression. Sensors 13(8):10027–10051. https://doi.org/10.3390/s130810027

Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J (2007) Nondestructive measurement of fruit and vegetable quality by means of nir spectroscopy: a review. Postharvest Biol Technol 46(2):99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024

Oliveira-Folador G, de Oliveira Bicudo M, de Andrade EF, Renard CMGC, Bureau S, de Castilhos F (2018) Quality traits prediction of the passion fruit pulp using NIR and MIR spectroscopy. LWT-Food Sci Technol 95:172–178. https://doi.org/10.1016/j.lwt.2018.04.078

Pan L, Zhang W, Zhu N, Mao S, Tu K (2014) Early detection and classification of pathogenic fungal disease in post-harvest strawberry fruit by electronic nose and gas chromatography–mass spectrometry. Food Res Int 62(8):162–168. https://doi.org/10.1016/j.foodres.2014.02.020

Qin J, Chao K, Kim MS, Lu R, Burks TF (2013) Hyperspectral and multispectral imaging for evaluating food safety and quality. J Food Eng 118(2):157–171. https://doi.org/10.1016/j.jfoodeng.2013.04.001

Siedliska A, Baranowski P, Zubik M, Mazurek W, Sosnowska B (2018) Detection of fungal infections in strawberry fruit by vnir/swir hyperspectral imaging. Postharvest Biol Technol 139:115–126. https://doi.org/10.1016/j.postharvbio.2018.01.018

Sturm K, Koron D, Stampar F (2003) The composition of fruit of different strawberry varieties depending on maturity stage. Food Chem 83(3):417–422. https://doi.org/10.1016/S0308-8146(03)00124-9

Sun Y, Wang Y, Xiao H, Gu X, Pan L, Tu K (2017a) Hyperspectral imaging detection of decayed honey peaches based on their chlorophyll content. Food Chem 235:194–202. https://doi.org/10.1016/j.foodchem.2017.05.064

Sun M, Zhang D, Li L, Wang Z (2017b) How to predict the sugariness and hardness of melons: a near-infrared hyperspectral imaging method. Food Chem 218:413–421. https://doi.org/10.1016/j.foodchem.2016.09.023

Thornton CR, Slaughter DC, Davis RM (2010) Detection of the sour-rot pathogen Geotrichum candidum in tomato fruit and juice by using a highly specific monoclonal antibody-based ELISA. Int J Food Microbiol 143(3):166–172. https://doi.org/10.1016/j.ijfoodmicro.2010.08.012

Williams, P. C. (1987). Qualitative applications of near infrared reflectance spectroscopy. Nearinfrared Technology in the Agriculture and Food Industries. Environment control in biology 42(3):217–223

Williamson B, Tudzynski B, Tudzynski P, Kan JALV (2007) Botrytis cinerea: the cause of grey mould disease. Mol Plant Pathol 8(5): 561–580. https://doi.org/10.1111/j.1364-3703.2007.00417.x

Wu D, Sun DW (2013) Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review—part I: fundamentals. Innov Food Sci Emerg Technol 19:1–14

Yu X, Yuan F, Fu X, Zhu D (2016) Profiling and relationship of water-soluble sugar and protein compositions in soybean seeds. Food Chem 196:776–782. https://doi.org/10.1016/j.foodchem.2015.09.092

Zhang L (2017). The effect of genes for sucrose metabolism and transportation on fruit sugar accumulation in strawberry (*Fragaria* × *ananassa Duch*). Gansu Agricultural University, p 31-52 (In chinese)