Food Chemistry 152 (2014) 316–322

Contents lists available at ScienceDirect

Food Chemistry

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

Analytical Methods

Effects of grown origin, genotype, harvest year, and their interactions of wheat kernels on near infrared spectral fingerprints for geographical traceability

Haiyan Zhao [a,b], Boli Guo [a], Yimin Wei [a,*], Bo Zhang [a]

[a] Institute of Agro-Products Processing Science and Technology, Chinese Academy of Agricultural Sciences/ Comprehensive Key Laboratory of Agro-Products Processing, Ministry of Agriculture, P.O. Box 5109, Beijing 100193, People's Republic of China

[b] College of Food Science and Engineering, Qingdao Agricultural University, Qingdao 266109, People's Republic of China

ARTICLE INFO

Article history:
Received 7 January 2013
Received in revised form 7 November 2013
Accepted 21 November 2013
Available online 27 November 2013

Keywords:
Wheat kernel
NIR
Geographical origin
Genotype
Harvest year
Traceability

ABSTRACT

The effects of origin, genotype, harvest year, and their interactions on wheat near infrared (NIR) spectra were studied to find the reasons for differences in NIR fingerprints of wheat from different geographical origins and the stability of NIR fingerprints among different years. Ten varieties were grown in three regions of China for 2 years. 180 kernel samples were analysed by NIR. The spectra after pre-treatment were analysed by principal component analysis, multi-way analysis of variance, and discriminant partial least-squares. The results showed that origin, genotype, year, and their interactions all had significant effects on wheat NIR fingerprints. The second overtones of N–H and C–H stretching vibrations and a combination of stretch and deformation of C–H group in wheat were mainly influenced by the geographical origin. The wavelength ranges 975–990 nm, 1200 nm, and 1355–1380 nm contained plenty of origin information to build robust discriminant models of wheat geographical origin.

1. Introduction

Agricultural products from different geographical origins have their own characteristics due to different climate, soil, and other environmental conditions (Rharrabti, Royo, Villegas, Aparicio, & García del Moral, 2003; Rharrabti, Villegas, Royo, Martos-Núñez, & García del Moral, 2003; Triboi et al., 2000). Many countries have published relevant regulations or laws to register and certify geographical indication brands of agricultural products, such as Regulation (EC) No 510/2006 and Measures for the Administration of Geographical Indications of Agricultural Products enacted by the Ministry of Agriculture of the People's Republic of China. However, some unscrupulous traders replace special local products with inferior or counterfeit products for economic profit, which may lead to an unfair competition in the agricultural product industry and harm the rights of consumers. Therefore, scientific identifying methods are needed as scientific support for the healthy development of special local product industry.

Among numerous analytical methods applied to food authentication in the last decade, the modern NIR analytical technique combined with chemometrics has the advantages of simplicity, speed, high-level efficiency, low cost, and being non-destructive. The NIR region spans the wavelength range of 780–2500 nm, and contains information about the relative proportions of C–H, N–H, and O–H bonds which are the primary structural components of organic molecules (Murray & Williams, 1987). NIR thereby builds a characteristic spectrum that behaves as a fingerprint of the sample. It has been increasingly adopted as an analytical tool for the identification and/or authentication of various products, by collecting samples randomly from different origins, such as virgin olive oil (Casale, Casolino, Oliveri, & Forina, 2010; Galtier et al., 2007; Woodcock, Downey & O'Donnell, 2008), tea (Chen, Zhao, & Lin, 2009; He et al., 2012), lotus root powder (Xu et al., 2012), honey (Woodcock, Downey & O'Donnell, 2009), and so on. These studies have demonstrated a high degree of discriminant success. Genotype is one of the important factors which affect chemical compositions of agricultural products (Ames, Clarke, Marchylo, Dexter, & Woods, 1999; Kindred et al., 2008; Wickramasinghe, Miura, Yamauchi, & Noda, 2005). Due to the adaptability of agricultural products to climate, main varieties may differ in their composition in different regions, which could lead to the change of NIR spectral fingerprints. In addition, the climate conditions (precipitation, temperature, sunshine time, etc.) in different years may be different in the same region, which results in the change of chemical compositions of agricultural products (Bontempo et al., 2009; Dornez et al., 2008; Kučerová, 2005). Therefore, NIR spectral

* Corresponding author. Tel.: +86 010 62815956; fax: +86 010 62895141.
E-mail addresses: xinyuyuanyin@163.com (H. Zhao), guoboli2007@126.com (B. Guo), weiyimin36@hotmail.com (Y. Wei), zjzb1978@126.com (B. Zhan.

0308-8146/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodchem.2013.11.122

fingerprints of agricultural products from the same region but different year may be different. The issues that the differences in NIR spectral fingerprints from different origins are caused by the difference of origin or genotype and whether characteristic fingerprints will change with varying genotypes and harvest years in the same origin are not yet clear. Therefore, it is necessary to study the effects of grown origin, genotype, harvest year, and their interactions on NIR spectral fingerprints to provide theoretical support for the application of NIR fingerprinting technique to identify the geographical origin of agricultural products. To our knowledge, no literature has been published on this aspect.

Wheat is a staple crop produced in China and in the world, and its distribution area is wide and adaptability is strong. Thus, it is easier to collect a large number of samples with the same genotypes in different origins and years and different genotypes in the same origin to study the effects of grown origin, genotype, harvest year, and their interactions on NIR fingerprints of agricultural products. 240 wheat samples were randomly collected during the 2008 and 2009 harvest periods from Hebei, Henan, Shandong, and Shaanxi provinces in China without considering the effect of genotype, and analysed by NIR analytical technique in our previous study. It was found that NIR spectral fingerprints were significantly different among wheat kernels from the same year but different origins. There were also some differences in NIR spectral fingerprints of wheat kernels from the same region but different years (Zhao, Guo, Wei, & Zhang, 2013). However, it has not been clear that the differences of spectral fingerprints among different regions result from the influence of origin or genotype and which fingerprints are relatively stable in the same region but different years.

The main objective of this study was to investigate the effects of grown origin, genotype, harvest year, and their interactions on wheat NIR spectral fingerprints based on the field experiment, and then examine the performance of identifying wheat geographical origin using the wavelengths associated with origin.

2. Materials and methods

2.1. Experimental design

Field plot experiments were conducted in three agricultural experiment stations of Zhaoxian (Hebei province), Huixian (Henan province), and Yangling (Shaanxi province) in China for 2 years (October 2010–June 2011 and October 2011–June 2012), respectively. Ten varieties (Han 6172, Heng 5229, Hengguan 35, Xinong 889, Xinong 979, Xiaoyan 22, Xinmai 18, Zhengmai 366, Zhoumai 16, and Zhoumai 18) were planted in each station each year. The size of every plot (variety) in each station was 10 m^2. The experimental design was a completely randomised block for a total of 10 plots in each station with no replication. Within each station, agricultural practices adopted the locally recommended wheat management, including seeding rates, fertilisation, irrigation, and the chemical control of weeds, pests and diseases.

2.2. Sampling

Wheat samples were collected randomly at three different sites in each plot during the harvest periods (June 2011 and June 2012). A 1 m^2 (1 m $\times$ 1 m) quadrant was placed randomly at each site for sampling, and the samples were collected at the same time by hand-cutting from the marked quadrant area. The whole-wheat samples were subsequently threshed, and the resulting kernel samples were retained for analysis. For each sample 100 g was chosen as an analytical sample. 30 samples were collected from each region each year.

The locations of each sampling station were recorded using a portable global positioning system receiver. Meteorological data were collected by automatic weather station. The information recorded related to these three regions is shown in Table 1.

2.3. Sample preparation

The kernel samples were rinsed with distilled water repeatedly to clear dust after picking out stones, weeds, etc., and then dried in an oven (DHG-9140A, Yiheng, Shanghai, China) at 38 °C for 10 h to make the moisture consistent as much as possible.

2.4. NIR spectroscopy measurement

Diode Array 7200 NIR Spectrometer (Perten Instrument AB, Sweden) was used to collect reflectance spectra of wheat kernels in the 950–1650 nm range and with a 5 nm resolution. The instrument was equipped with a circle sample cup (75 mm in diameter and 25 mm in depth) rotating on itself during the measurement. Each sample was measured four times and the mean spectrum was used in further data analysis. NIR spectra were collected using DA 7200 Simplicity software (Perten Instrument AB, Sweden), stored in absorbance format, and then exported for chemometric analysis.

2.5. Spectral data pre-treatment

It is necessary to perform mathematical pre-treatment to reduce the systematic noise, such as baseline variation, light scattering, path length differences, and so on. Raw spectra were exported from WINISI in JCAMP.DX format and imported into the Unscrambler (version 9.7, CAMO ASA, Norway) for data pre-treatment using standard normal variate (SNV) followed by the second derivative. SNV was applied to correct light scatter and reduce the changes of light path length. The second derivative was performed using Savitzky–Golay method (11 point smoothing and third-order polynomial) to eliminate baseline variation and enhance spectral features.

2.6. Chemometric methods

2.6.1. Principal component analysis (PCA)

PCA is a method of data reduction that constructs new uncorrelated variables, known as principal components (PCs) that are linear combination of the original ones. PCA was used to reduce the dimensionality of spectral data of wheat samples while retaining as much information as necessary in this study.

2.6.2. Multi-way analysis of variance (Multi-way ANOVA)

Multi-way ANOVA provides analysis of variance for multiple factor variables by one or more dependent variables. The factor variables divide the population into groups. Using this procedure, null hypotheses about the effects of factor variables on the means of various groupings of a joint distribution of dependent variables could be tested. Origin, genotype, year, and their interactions were factor variables and PC scores were dependent variables in this study. Multi-way ANOVA using Wilks' lambda criterion with approximate F statistic was used to test the influence of origin, genotype, year, and their interactions on wheat NIR spectra as well as the univariate analysis of variance for each dependent variable.

2.6.3. Discriminant partial least squares (DPLS)

DPLS was applied in this work to classify wheat from different regions. Due to the fact that we had more than two variable groups, PLS2 algorithm was used to develop classification models. PLS2 regression used the dummy variable as the dependent variable Y