Food Chemistry 227 (2017) 342–348

Contents lists available at ScienceDirect

Food Chemistry

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

Development of predictive models for total phenolics and free *p*-coumaric acid contents in barley grain by near-infrared spectroscopy

Zhigang Han, Shengguan Cai, Xuelei Zhang, Qiufeng Qian, Yuqing Huang, Fei Dai, Guoping Zhang *

Department of Agronomy, Zhejiang Key Laboratory of Crop Germplasm, Zhejiang University, Hangzhou 310058, China

ARTICLE INFO

Article history:
Received 21 October 2015
Received in revised form 1 January 2017
Accepted 13 January 2017
Available online 16 January 2017

Chemical compounds studied in this article:
p-Coumaric acid (PubChem CID 637542)
Ferulic acid (PubChem CID 445858)
Gallic acid (PubChem CID 370)
Formic acid (PubChem CID 284)

Keywords:
Barley (*Hordeum vulgare* L.)
Phenolic compound
Near infrared spectroscopy (NIRS)
Partial least squares (PLS)
Least squares support vector machine (LS-SVM)

ABSTRACT

Barley grains are rich in phenolic compounds, which are associated with reduced risk of chronic diseases. Development of barley cultivars with high phenolic acid content has become one of the main objectives in breeding programs. A rapid and accurate method for measuring phenolic compounds would be helpful for crop breeding. We developed predictive models for both total phenolics (TPC) and *p*-coumaric acid (PA), based on near-infrared spectroscopy (NIRS) analysis. Regressions of partial least squares (PLS) and least squares support vector machine (LS-SVM) were compared for improving the models, and Monte Carlo-Uninformative Variable Elimination (MC-UVE) was applied to select informative wavelengths. The optimal calibration models generated high coefficients of correlation (r_{pre}) and ratio performance deviation (RPD) for TPC and PA. These results indicated the models are suitable for rapid determination of phenolic compounds in barley grains.

1. Introduction

Barley (*Hordeum vulgare* L.) is the fourth largest cereal crop worldwide in terms of planting area, and used mainly for feed and beer production (Cai et al., 2015; Goupy, Hugues, Boivin, & Amiot, 1999; Madhujith & Shahidi, 2009). Recently, more attention has been given to processing functional food from barley, as it is rich in phenolic compounds, which are associated with reduced risk of chronic inflammation, cardiovascular diseases, cancers, and diabetes (Chandrasekara & Shahidi, 2010; Slavin, 2003). Phenolic acids in barley grains exist in free, bound and soluble forms, and consist mainly of ferulic and *p*-coumaric acids (Kim et al., 2007).

High phenolic acid content is regarded as an important objective in barley breeding. Rapid and accurate measurement of phenolic acid content is a prerequisite for efficient identification of elite germplasm and breeding lines. However, the conventional method, i.e. chemical analysis of phenolic acid, is time-consuming and labor-intensive, limiting improvement of phenolic acid content in barley breeding. Obviously, development of a simple, rapid and effective method for measuring phenolic acid content is imperative.

At present, near-infrared reflectance spectroscopy (NIRS) has been employed widely as a high-throughput method for measurement of chemical components in quality evaluation of agricultural and food products (Cen & He, 2007). The technique is based on correlation between chemical properties and absorption of light at specific wavelengths in the near-infrared region (Moron & Cozzolino, 2002). Currently, NIRS is used intensively as an alternative method for inspection of food quality, such as amino acids content (Bao et al., 2012) and other nutritional values (Tarr, Diepeveen, & Appels, 2012) in barley, and biochemical quality parameters in cocoa (Krähmer et al., 2015). It has also been applied successfully for the determination of phenolics, such as phenols in sorghum grains (Dykes, Hoffmann, Portillo-Rodriguez, Rooney, & Rooney, 2014), phenolics and flavonoids in rice grains (Zhang, Shen, Chen, Xiao, & Bao, 2008), and phenolic compounds in green rooibos (Manley, Joubert, & Botha, 2006). However, there is no report describing the use of NIRS technique to measure phenolics in barley.

* Corresponding author.
E-mail address: zhanggp@zju.edu.cn (G. Zhang).

http://dx.doi.org/10.1016/j.foodchem.2017.01.063
0308-8146/

Accordingly, the current study aimed to determine whether phenolic acid content could be predicted by NIRS in barley grains, and develop calibration equations for the estimation of these components. In order to obtain robust and optimal calibration specificity, we used Monte Carlo-Uninformative Variable Elimination (MC-UVE) to select informative wavelengths. Here, the different calibration models established by partial least squares (PLS) with 10 pretreatments and least squares support vector machine (LS-SVM) are presented.

2. Plant materials and methods

2.1. Plant materials

In total, 90 Tibetan wild barley accessions (grown in 2013 and 2014) and 40 cultivated barley genotypes (grown in 2014) were planted at the Zijingang farm station of Zhejiang University (Hangzhou, China, 30°22′N, 119°26′E). The plot size was 0.75 m^2 (three rows with 3 m long and 0.25 m between rows). The experiment was arranged in a randomized complete block with three replicates. At maturity, barley grains were harvested, dried and then stored in a cool room (4 °C) for further analysis.

2.2. Total phenolics analysis

Grain samples were finely ground with a 0.5 mm screen. Phenolic compounds were extracted according to the procedure described by Zhao and Gu (2006) with minor modifications. Briefly, the sample (around 200 mg) was sonicated (40 kHz, 120 W) for 1 h with 4 ml of 80% methanol (v/v) at 25 °C. After centrifugation (10000*g*, 20 min), the supernatants were collected and stored at 4 °C. Total phenolic content (TPC) of the extracts was determined with three replicates according to the Folin-Ciocalteu spectrophotometric method (Zhao and Gu, 2006) and expressed as micrograms of gallic acid equivalents per gram of dry barley (μg of GAE/g of db). The average recovery (%), correlation coefficient (r^2), linear range (μg/ml) and relative standard deviation (RSD) were also calculated.

2.3. p-Coumaric acid analysis

The measurement of *p*-coumaric acid (PA) was based on the method described by Cai et al. (2015): 200 mg of samples was sonicated (40KHz, 120 W) for 40 min with 4 ml of 80% methanol (v/v) at 70 °C. After centrifugation (10000*g*, 20 min), the supernatant was evaporated to dryness in a vacuum freezing dryer. The residue was dissolved in 250 μL of 70% HPLC grade methanol (v/v) and filtered through a 0.45 μm membrane. Diamonsil 5u C18 (250 × 4.6 mm) column (Dikma, China) was used for the separation of phenolic acids at 40 °C. The mobile phase consisted of Solvent A (0.1% formic acid in water) and Solvent B (100% methanol). A gradient procedure was used for elution: 0 min, 30% B; 10 min, 45% B; 20 min, 50% B; 25 min, 80% B; 32 min, 30% B; 37 min, 30% B. PA was identified with three replicates based on their relative retention times. An external standard method with *p*-coumaric acid was used for quantification. For total phenolics, the correlation coefficient (r^2), linear range (μg/ml), average recovery (%), relative standard deviation (RSD) were calculated. In addition, standard and sample curves were recorded.

2.4. NIRS measurement

The ground barley samples (about 2.5 g) were scanned with NIRS mono-chromator and the corresponding spectra collected using a NIRSystems 5000 (Silver Spring, USA) instrument in reflectance mode. Each sample was scanned with four replicates in a circle sample cup (35 mm in diameter and 18 mm in depth). The spectrum was collected from 1100 to 2500 nm with an increment 2 nm using the software WinISI II (InfraSoft International, USA). The average spectrum of four recordings for each sample was used for further analysis. In modeling, the Kennard-Stone (KS) algorithm (Kennard & Stone, 1969) was performed for distributing all samples into calibration and prediction sets with a ratio of 3:1. The calibration set was used for calibration models, and the predictive capabilities and analytical features of the calibration models were validated using the prediction set.

2.5. Spectral analysis

In order to improve the performance of any spectral model, several spectral preprocessing methods including the Savitzky-Golay smoothing (SGS), area normalization, multiplicative scatter correction (MSC), first derivative (1st D) and standard normal variate (SNV), were implemented. The pre-treatments were carried out according to the instructions of Unscrambler V9.7 (CAMO PROCESS AS, Oslo, Norway).

The MC-UVE method was applied for selection of characteristic (informative) wavelengths (variables). Only sample- or component-specific information was retained (Li, 2012). MC-UVE can identify and encode more aspects of the relationship between independent and dependent variables (Cai, Li, & Shao, 2008). MC-UVE was performed using MATLAB (Version 7.8.0.347, The Math Works. Inc US). Partial least-squares regression (PLS) was analyzed to determine the relationship between a set of independent spectral variables (*X*) and a single dependent variable (*Y*). With the capability for both linear and non-linear multivariate calibration, LS-SVM can solve multivariate calibration problems relatively quickly (Wu et al., 2012). The standard LS-SVM algorithm was defined by Suykens and Vanderwalle (1999) and implemented based on the LS-SVM toolbox of MATLAB (Version 7.8.0.347, The Math Works. Inc US) to derive LS-SVM models.

2.6. Model evaluation

The performance of regression models was evaluated by standard error of calibration (SEC), standard error of prediction (SEP), and the correlation coefficient (r) between the predicted and measured parameters (Lin et al., 2014). In addition, the residual predictive deviation (RPD) was performed for evaluating the quality of regression models (Arana, Jarén, & Arazuri, 2005; Fearn, 2002). Generally, higher r_{cal}, r_{pre} and RPD values, and lower RMSEC and RMSEP values translate to more reliable prediction of chemical composition (Wu et al., 2012).

Table 1
Reference chemical data for TPC and PA contents of barley grains in calibration and validation sets.

	Calibration set				Validation set			
Constituent[a]	No.	Mean	SD	Range	No.	Mean	SD	Range
TPC	135	1914.79	373.34	887.74–2573.68	45	1800.98	444.36	934.30–2441.76
PA	165	0.85	0.44	0.25–1.90	55	0.74	0.44	0.24–1.75

Note: TPC: total phenolics; PA: *p*-coumaric acid; GAE: gallic acid equivalents. DB: dry barley; SD: standard Deviations.
[a] Phenolic content was expressed as μg GAE/g DB, *p*-coumaric acid was expressed as μg PA/g DB.