

# Determination of Main Categories of Components in Corn Steep Liquor by Near-Infrared Spectroscopy and Partial Least-Squares Regression

Xue Xiao,<sup>†,‡,§,||</sup> Yuanyuan Hou,<sup>†,‡,||</sup> Jun Du,<sup>#,⊗</sup> Yang Liu,<sup>†,‡</sup> Yanjie Liu,<sup>†,‡,⊥</sup> Linyi Dong,<sup>†,‡</sup> Qionglin Liang,<sup>§</sup> Yiming Wang,<sup>§</sup> Gang Bai,<sup>†,‡</sup> and Guoan Luo<sup>\*,†,§</sup>

<sup>†</sup>College of Pharmacy, Nankai University, Tianjin 300071, People's Republic of China

<sup>‡</sup>State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, People's Republic of China

<sup>§</sup>Department of Chemistry, Tsinghua University, Beijing 100084, People's Republic of China

<sup>#</sup>College of Life Science, Nankai University, Tianjin 300071, People's Republic of China

<sup>⊗</sup>China Biotech Fermentation Industry Association, Beijing 100833, People's Republic of China

<sup>⊥</sup>College of Pharmacy, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, People's Republic of China

**ABSTRACT:** Corn steep liquor (CSL) is an important raw material that has a high nutritional value and serves as a nitrogen source. This study aimed to develop a fast, versatile, cheap, and environmentally safe analytical method of quantifying the total acidity (TA) of CSL as well as its contents of dry matter (DM), total sugars (TS), total reducing sugars (TRS), total free amino acids (TFAA), total nitrogen (TN), and total sulfite (TSu). The near-infrared (NIR) spectroscopy measurements of 66 samples (22 batches) of CSL were analyzed by partial least-squares regression using several spectral preprocessing methods. Multivariate models developed in the NIR area showed good predictive abilities for DM, TA, TS, TRS, TFAA, TN, and TSu determination. These results confirm the feasibility of the multivariate spectroscopic approach as a replacement for expensive and time-consuming conventional chemical methods. Thus, a convenient and feasible method for the quality control of fermentation raw materials for food additives and fine chemicals, especially in CSL, is established.

**KEYWORDS:** corn steep liquor, NIR, amino acid, PLSR, PCA

## INTRODUCTION

Biomanufacturing technology is rapidly developing. However, uncertainties in the quality control of raw materials lead to unstable fermentation.<sup>1,2</sup> Recent studies report the development of byproducts using advanced process analytical technology (PAT) to discover a highly feasible method that increases the production yield and sugar–acid transfer ratio.<sup>3–6</sup>

Corn is used to produce corn starch. The main byproduct and surplus raw material is corn steep liquor (CSL). The ratio of the consumption of food additives and fine chemicals from corn to the total consumption is about 30%, about 50 million tons per annum, over the past 3 years.<sup>7</sup> Controlling the quality of the corn raw materials to improve the utilization efficiency means saving corn resources. CSL is a soluble solid that constitutes nearly 40–50% (w/w) of the dry weight of corn and consists of a mixture of reducing sugars and amino acids (mostly free amino acids).<sup>8,9</sup> Previous results of PAT methods used to describe the acidity of CSL as well as its contents of total sugars (TS), total reducing sugars (TRS), total free amino acids (TFAA), total nitrogen (TN), and total sulfite (TSu) indicate the importance of CSL as a substrate material for fermentation purposes, even though its physical and chemical functions are influenced by the corn variety.<sup>10</sup> CSL can be used as a substrate material for the production of glutamic acid, penicillin, lactic acid, hyaluronic acid, and others. It has been studied for its potential use as a nutritional and functional

supplement to water-soluble plant proteins and vitamins in the fermentation process.<sup>11–13</sup>

The use of CSL as a biotechnological substrate, nitrogen source, and carbon source largely depends on its chemical composition.<sup>14</sup> At the beginning of fermentation, the ready-made carbon frames of CSL are consumed by bacteria. The quality of CSL and its organic nitrogen content are important in the fermentation process.<sup>15–17</sup> However, the chemical characterization and quantification of dry matter (DM), total acidity (TA), TS, TRS, TFAA, TN, and TSu by traditional analyses require tedious sample preparation and pose inherent environmental risks, in addition to being time-consuming and expensive. Nevertheless, TRS and TFAA are important in the entire fermentation process.

Several types of food have been analyzed using near-infrared spectroscopy (NIR), and a number of multivariate calibrations have been performed mainly using partial least-squares (PLS) regression (PLSR).<sup>18–23</sup> However, no research on CSL using both NIR and PLSR has been reported. In the current work, multivariate models were developed to quantify DM, TA, TS, TRS, TFAA, TN, and TSu in CSL using NIR and PLSR. CSL was classified using NIR and principal component analysis

**Received:** March 27, 2012

**Revised:** July 29, 2012

**Accepted:** July 29, 2012

**Published:** July 30, 2012

(PCA). The current study provides a simple and rapid method for the determination of the main categories of CSL components by NIR-PLSR.

## MATERIALS AND METHODS

**CSL Samples.** Twenty-two batches (three samples for each batch) of CSL consisting of commercial and experimental varieties were used to build the regression models. These samples were collected from the provinces of Shandong (five batches), Henan (eight batches), and Hebei (six batches). The commercial samples were bought from a local retail market. All samples were refrigerated at 4 °C and stored at room temperature before measurements.

**Wet Chemistry Analysis.** To determine DM, a halogen moisture analyzer (model HB43-S, Mettler Toledo, Swiss Confederation) was used at mode A under an auto condition. The sample weights ranged from 2.6 to 3.5 g per specimen. TA was determined by titrimetry and expressed as weight percent hydrochloric acid. A 0.01 mol/L NaOH solution was used for the titration, and phenolphthalein was used to indicate the pH end point of the reaction. TSu was also determined by titrimetry and expressed as weight percent hydrochloric acid. An iodine titration solution was used to determine the TSu concentration of the remaining solution after TA determination, and starch was used as an indicator of the end point of the reaction.

The corresponding colorimetries were independently used to determine the TS and TRS concentrations. TS quantification was performed using a phenol–sulfuric acid method, and the absorbance at 488 nm was recorded using a Shimadzu spectrophotometer (model 1800; Kyoto, Japan) and a quartz cuvette with a path length of 1 cm. TRS was detected by the 3,5-dinitrosalicylic acid method.<sup>24</sup> The solute fraction was diluted to adjust the concentration according to the sensitivity of the analytical technique (from 12.0 to 200.0 µg/mL for TS and TRS) prior to the analysis.

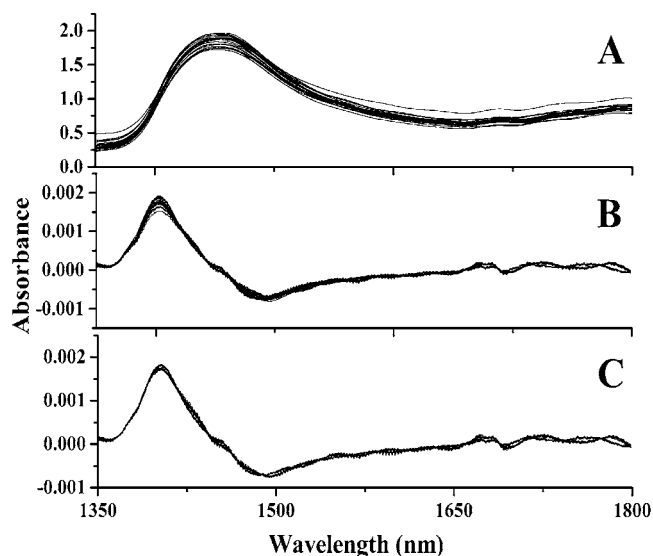
The determination of TFAA, expressed as milligrams per milliliter of free amino acids, was performed by high-performance liquid chromatography (HPLC).<sup>25</sup> More than 10 free amino acids are important CSL components. An ATN-100 Kjeldahl nitrogen analyzer (Hongji Instrument Co., Ltd., Shanghai, China) comprising a digestion furnace with a pit furnace and an azotometer distiller was used to determine TN.

The sample (0.2 g) was digested with potassium sulfate (5 g), copper sulfate (0.3 g), and concentrated sulfuric acid (10 mL) at 450 °C for 25 min until the solution turned green. After 10 min of distillation with certain amounts of NaOH and H<sub>2</sub>O, sulfuric acid (0.05 mol/L) was used as a titrant to determine the end point; that is, the receiving bluish green liquid becomes gray-purple.

All analytical determinations were conducted in triplicate; each sample was filtered using a 0.45 µm Millipore membrane prior to the determination.

**Spectroscopic Measurements and Data Pretreatment.** NIR spectra were obtained from 1350 to 1800 nm at an interval of 0.1 nm and at a total of 32 scans using an Axsun NIR Analyzer XL 410 spectrophotometer (Axsun Technologies Inc., Billerica, MA, USA). A total of 66 spectra were obtained for all samples and expressed as reflectance values (percent). The spectra were pretreated to remove irrelevant information and noise before the model was calibrated. Preprocessing methods such as nonpretreatment, convolution smoothing, one-dimension convolution (one-DC), two-dimension convolution (two-DC), multiple-scatter correction (MSC), standard normal variable transformation, and spectral data normalization were used to deal with the CSL data. The typical and selected pretreated spectra are shown in Figure 1.

**PLS and Latent Variable (LV) Selection.** The spectral and concentration data were organized in matrices, and the multivariate calibration model was developed using THUNIR (3rd ed, Tsinghua University, Beijing, China).<sup>26–29</sup> The number of LVs, namely, the main factors, was selected on the basis of the minimization of the root-mean-square error of cross-validation (RMSECV), which corresponds to the predictive error obtained in the cross-validation stage. The step is similar to the cross-validation, wherein a parameter based on the



**Figure 1.** Typical raw spectra (A) and pretreated spectra: (B) one-dimension convolution spectra; (C) one-dimension convolution plus multiple scatter correction spectra of corn steep liquor samples (total of 66 samples).

division of the calibration is set into subgroups, and each subgroup is sequentially removed for inclusion in the predictive set.<sup>30</sup>

**PCA.** A PCA of the NIR spectra was conducted for the quality control analysis of CSL.<sup>31–33</sup> The main purpose of PCA is dimension reduction to obtain a few new variables that can express the data characteristics of the original variables without loss of information as much as possible. All data were fed into the PCA. All programs were operated using Unscrambler V 9.7 (Camo Software AS, Oslo, Norway).

## RESULTS AND DISCUSSION

**Conventional Chemical Analysis.** Table 1 shows the average results for the measured values of DM (% w/w), TA (% w/w), TS (mg/mL), TRS (mg/mL), TFAA (mg/mL), TN (% w/w), and TSu (% w/w), as well as the standard deviation (SD) from 66 selected CSL samples. TFAA had the lowest concentration and a large coefficient of variation (CV) from selected CSL samples (16%). These results closely agree with the values from previous studies.<sup>34,35</sup> DM had the highest average concentration (37.83%) and the lowest CV (9%), whereas TN and TSu had CVs (12%) that were significantly lower than those previously reported. TRS (CV = 59%) had the highest sample heterogeneity followed by TS (CV = 43%). The average acidity was 9.75%, and a low CV of 13% was observed.

Although the SD of DM was small, the SDs of TS and TFAA were large and different from each other, significantly affecting the quality of CSL as a fermentation substrate. The credibility and repeatability of the conventional analytical methods are high but time-consuming.

**NIR Analysis.** The NIR spectra of CSL were acquired from 1350 to 1800 nm. The current study mainly used two pretreatment methods. One is DC, which effectively eliminates the interference of the baseline and background and distinguishes overlapping peaks to enhance the resolution and sensitivity. The other is MSC, which eliminates the scattering effects produced by the inhomogeneous distribution and irregular form of the particles. MSC is used to analyze the diffuse reflection of solid samples and transmission reflection of slurry substances.

**Table 1. Determination of Dry Matter (DM), Total Acidity (TA), Total Sugars (TS), Total Reducing Sugars (TRS), Total Free Amino Acids (TFAA), Total Nitrogen (TN), and Total Sulfite (TSu) in Corn Steep Liquor by Wet Chemical Methods**

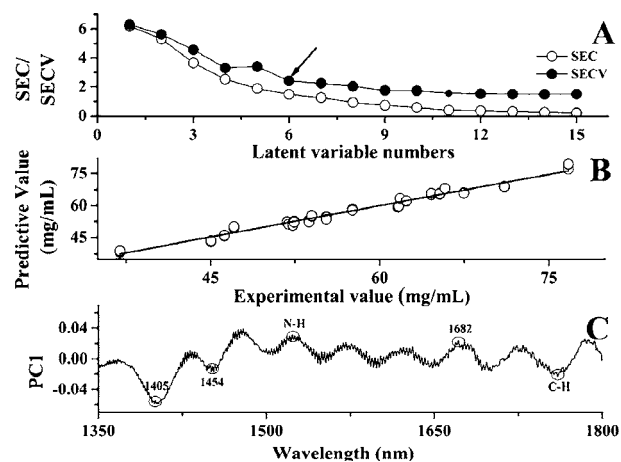
batch no.	code <sup>a</sup>	DM % g/g	TA (% w/w)	TRS (mg/mL)	TS (mg/mL)	TFAA (mg/mL)	TN (% w/w)	TSu (% w/w)
01003	1-C; 2-C; 3-C	40.58	9.23	11.04	48.71	76.71	47.26	0.0141
01004	4-C; 1-V; 5-C	42.65	8.79	5.41	25.77	65.78	53.60	0.0148
01006	6-C; 7-C; 8-C	39.19	8.94	19.31	78.86	62.35	45.24	0.0146
01015	9-C; 10-C; 11-C	37.26	12.34	4.28	18.81	64.56	50.21	0.0141
01017	12-C; 13-C; 2-V	40.25	12.22	2.74	19.20	71.05	55.82	0.0151
01018	3-V; 4-V; 14-C	40.55	11.23	3.13	18.85	52.42	46.37	0.0146
01019	15-C; 16-C; 17-C	42.29	10.77	3.30	18.71	61.68	46.24	0.0151
01020	18-C; 19-C; 20-C	44.08	9.23	3.92	18.88	61.60	46.51	0.0147
02001	5-V; 21-C; 22-C	40.72	8.86	4.73	31.20	61.81	47.73	0.0150
02002	23-C; 24-C; 25-C	35.39	10.54	5.22	27.70	51.90	67.72	0.0118
02003	26-C; 27-C; 28-C	37.19	9.65	7.41	39.24	46.23	50.34	0.0122
02004	29-C; 30-C; 31-C	37.76	11.46	7.79	39.58	36.97	60.77	0.0150
02005	32-C; 6-V; 7-V	39.05	9.34	8.71	45.90	53.97	46.80	0.0076
02006	33-C; 34-C; 35-C	38.59	10.52	6.50	36.08	45.04	48.01	0.0120
02007	8-V; 36-C; 9-V	37.39	8.75	2.03	17.04	47.07	48.22	0.0147
02008	37-C; 38-C; 39-C	35.09	7.70	10.55	45.35	57.58	48.41	0.0140
02009	40-C; 41-C; 42-C	33.16	9.33	6.04	42.06	53.74	49.60	0.0143
02010	43-C; 44-C; 45-C	33.95	8.86	4.98	34.14	51.80	52.65	0.0147
02011	46-C; 47-C; 10-V	38.13	8.39	6.79	39.96	67.47	48.35	0.0133
02012	48-C; 49-C; 50-C	31.74	9.83	11.84	45.97	65.33	61.73	0.0136
02013	51-C; 52-C; 53-C	33.05	9.45	4.53	21.63	52.30	59.55	0.0144
02014	54-C; 55-C; 56-C	34.12	9.07	5.43	30.64	55.26	53.93	0.0143
min		31.74	12.34	19.31	78.86	76.71	67.72	0.0151
max		44.08	12.34	19.31	78.86	76.71	67.72	0.0151
av		37.83	9.75	6.62	33.83	57.39	51.593	0.0138
SD <sup>b</sup> (C)		3.44	1.19	3.91	14.62	9.41	6.20	0.0016
SD (V)		1.68	1.35	2.58	11.95	6.65	3.23	0.0030
CV <sup>c</sup> (%)		9	13	59	43	16	12	12

<sup>a</sup>Code of the sample numbers of calibration (C) and external validation (V). <sup>b</sup>SD, standard deviation. <sup>c</sup>CV, coefficient of variation [(standard deviation/average) × 100].

The bands corresponding to the functional groups of CSL were extremely difficult to identify because of the complicated and overlapping signals. Thus, the automatic optimization software THUNIR was used to obtain the best model that had the lowest LV and RMSECV. THUNIR is a dedicated chemometrics software created by Tsinghua University. It is used in several pretreatment and band-selection methods. Its core algorithm is PLS.

**PLS Models.** The sample selections for the calibration and external validation used compatible and typical sets of the determination analysis as key factors. Fifty-six spectra were used for the calibration group, whereas 10 other spectra were used as an external validation set, representing the seven varieties of CSL from different firms (Table 1). The SD values showed that the calibration and external validation sets were similar to the contents. The models were built using the NIR regions, several LV numbers, and a number of signal pretreatment methods to achieve high predictive ability. The best models were selected using suitable and smaller square error of cross-validation (SECV) values. The selection process for LV and SECV was automated using THUNIR.

Figure 2A shows that the introduction of the sixth LV minimized the RMSECV (SECV) value. Under this condition, the model exhibited good predictive performance, providing high coefficient values for the prediction of the calibration set (Figure 2B) and lower error values in the cross-validation (SECV of 1.5003). Figure 2C shows the X-loading score of



**Figure 2.** Square error of cross-validation (SECV)/standard error of calibration (SEC) versus latent variable number (LV): (A) introduction of the sixth LV minimized the RMSECV (SECV) value [high coefficient values between predictive values and experimental values for the calibration set (56 samples)]; (B) X-loading score of the model for total free amino acids (TFAA) for the raw spectra; (C) good possibility to predict TFAA.

TFAA for the raw spectra, which were used in the TFAA model. Figure 2C indicates the good possibility to predict the content. The TFAA absorption signals were obtained from 1600 to 1800 nm. The first main overtone absorption band

**Table 2. Parameters of Calibration Models of Dry Matter (DM), Total Acidity (TA), Total Sugars (TS), Total Reducing Sugars (TRS), Total Free Amino Acids (TFFA), Total Nitrogen (TN), and Total Sulfite (TSu)**

model	LV <sup>a</sup>	SPNIRS <sup>b</sup>	bands (nm)	SEC <sup>c</sup>	SECV <sup>d</sup>	fitting equation	CD <sup>e</sup>	offset
DM	8	one-DC <sup>f</sup> 11,3	1619–1799	0.4966	0.8517	$y = 0.9798x + 0.7572$	97.98	-0.0012
TA	7	one-DC 11,3	1619–1799	0.2751	0.4443	$y = 0.9484x + 0.5040$	94.84	-0.0031
TS	8	one-DC 11,3	1619–1709	2.1554	3.5203	$y = 0.9788x + 0.7381$	97.88	-0.0471
TRS	9	one-DC 11,3 + MSC <sup>g</sup>	1529–1709	0.9290	0.3668	$y = 0.9915x + 0.0594$	99.15	-0.1967
TFAA	6	one-DC 11,3	1619–1799	2.4146	1.5003	$y = 0.9766x + 1.3436$	97.66	0.0634
TN	8	one-DC 11,3	1529–1799	2.7426	1.0973	$y = 0.9694x + 1.5934$	96.94	-0.3934
TSu	5	one-DC 11,3	1350–1439 1529–1799	0.0009	0.0004	$y = 0.9219x + 0.0011$	92.19	0.0000

<sup>a</sup>Latent variable number. <sup>b</sup>Spectrum pretreatment in near-infrared spectroscopy. <sup>c</sup>Standard error of calibration. <sup>d</sup>Standard error of cross-validation. <sup>e</sup>Coefficient of determination. <sup>f</sup>Dimension convolution. <sup>g</sup>Multiple scatter correction.

corresponding to N–H of the amino acids was observed between 1530 and 1600 nm. The first overtone C–H absorption of methyl was in the bands between 1700 and 1800 nm. The C–H combination bands appeared within the range of 1300–1500 nm. A C–H peak corresponding to the carbonyl functional group separated by a saturated carbon appeared at 1682 nm.<sup>36</sup> An absorption peak near 1454 nm (around 1445 nm) may be the first overtone corresponding to free OH and/or carboxylic acid.<sup>37</sup>

The regression model with the best prediction for the TFAA concentration was obtained by one-dimension convolution on the spectra and PLS, which was selected from the models using pretreatment methods such as nonpretreatment, one-DC, and one-DC+MSC for the spectral data. The true and predicted values indicate that the model had a good predictive capacity (Figure 2B). The correlation coefficients showed that the spectral bands of highest importance for the established correlation were those corresponding to the COO<sup>-</sup> and NH<sub>4</sub><sup>+</sup> groups of the amino acids.

The determination of DM by the one-DC of the NIR region produced favorable effects. The model used a number of LVs (Table 2) to gain a small RMSECV, resulting in a more accurate prediction. The best predictive capacity for TS was acquired using the one-DC of the spectral data within 1619–1709 nm with eight LVs. The correlation between the real and predicted values was interesting. The TRS model was acquired through a one-DC and multiple-scatter correction of the spectral bands from 1529 to 1709 nm with nine LVs. Similar to the TS model, the predictive accuracy of TRS was not as good as that of the other models. This result may be due to the instability of the samples.

The best outcome and lowest predictive errors for the TA concentration in CSL were obtained with the one-DC of the spectral data from 1619 to 1799 nm with seven LVs. The best predictive capacities of all models are shown in Table 3, which also shows the experimental and predictive values of all compounds in CSL (sample set for external validation). Under normal circumstances, the multivariate calibration models confirmed the predictive capacities represented by the lower relative standard errors. The models used in the present study generally showed better predictive capacities than other similar methods and used lower and more appropriate numbers of latent variables.

The prediction of the DM, TA, TS, TRS, TFAA, TN, and TSu concentrations by the regression models showed a lower standard error for the external validation sample set compared with that of the conventional methodologies (Table 3). Except for TN and DM, the ratios of the standard deviation of the

validation set to the standard error of prediction were >3, illustrating the feasibility to determine the main categories of CSL components by NIR and PLSR.

**PCA Result.** NIR spectroscopy is a widely used efficient technique for the quality control of complex mixtures, especially foods and wines. Multivariate data analysis is often combined with NIR spectroscopy, and the data are pretreated using MSC, one-DC, and other techniques. The homogeneity of CSL is the quantitative basis of composition determination by NIR. The heterogeneity of CSL may be due to the different production processes and causes different raw materials to cluster initially. PCA is useful in distinguishing and efficiently separating raw materials that come from different commercial batches (Figure 3). The samples from different companies were clustered. The homogeneity of the samples from the same manufacturer was better than that of the samples from different manufacturers. One batch sample varied from the other samples of the same manufacturer, and a batch of commercial samples was close to the manufacturer in Shandong, suggesting that the differences between samples of the same manufacturer were still large. The classification results indicated that the PCA method can be used in the quality control of CSL by clustering, which enables easy distinction of the CSL source. The results suggest that all materials prepared using different pretreatment processes, manufacturing procedures, and storage conditions vary in quality. All of these factors can lead to significant differences in CSL quality. Therefore, the rapid and accurate determination of CSL is very convenient and allows fermentation to proceed smoothly.

This study was undertaken to establish the feasibility of using NIR spectroscopy to predict the main components of CSL, including DM, TA, TS, TRS, TFAA, TN, and TSu. The seven PLS models built by their most appropriate wavebands and LVs showed good predictive capacities for the analysis of CSL samples. The average value of the prediction was equally accurate, indicating that the NIR-PLS method can suitably replace the laborious traditional chemical analyses. The rapid determination of the major CSL components, especially DM, TS, and TFAA, confirms the feasibility of using NIR-PLSR to realize accurate and systematic fermentation with the right concentrations.

The rapid and simultaneous quantification of CSL components is convenient for regulating the fermentation products. The appropriate adjustment of parameters depending on the variable quality of CSL makes the fermentation process stable and the level of lower power consumption controllable. The advantages of NIR spectroscopy and the PLS method are their abilities to develop multivariate regression models with

Table 3. Predictive Values versus Experimental Values of Dry Matter (DM), Total Acidity (TA), Total Sugars (TS), Total Reducing Sugars (TRS), Total Free Amino Acids (TFFA), Total Nitrogen (TN), and Total Sulfite (TSu) in the Validation Set

sample	DM		TA		TS		TRS		TFFA		TN		TSu		
	EV <sup>a</sup>	PV <sup>b</sup>	EV <sup>a</sup>	PV <sup>b</sup>	EV <sup>a</sup>	PV <sup>b</sup>	EV <sup>a</sup>	PV <sup>b</sup>	EV <sup>a</sup>	PV <sup>b</sup>	EV <sup>a</sup>	PV <sup>b</sup>	EV <sup>a</sup>	PV <sup>b</sup>	RD <sup>c</sup>
01004-2	42.65	42.44	8.79	9.00	25.77	28.24	5.41	5.59	45.04	43.85	53.60	53.24	0.0148	0.0149	0.84
01017-3	40.25	40.56	12.22	12.30	19.20	18.42	2.74	3.44	51.80	52.16	55.82	56.08	0.0151	0.0145	-4.10
01018-1	40.55	41.42	11.23	11.29	18.85	12.79	3.13	3.20	51.90	51.32	46.37	45.99	0.0146	0.0152	4.17
01018-2	40.55	41.25	11.23	11.37	18.85	10.36	3.13	2.81	51.90	51.33	46.37	47.54	0.0146	0.0154	5.43
02001-1	40.72	40.60	8.86	9.11	31.20	31.25	4.73	5.40	53.74	53.22	47.73	48.72	0.0150	0.0153	2.02
02005-2	39.05	37.67	9.34	9.93	45.90	46.27	8.71	9.37	61.60	59.57	46.80	50.45	0.0076	0.0092	20.76
02005-3	39.05	37.70	9.34	10.01	45.90	47.33	8.71	9.46	61.60	59.20	46.80	50.67	0.0076	0.0090	18.69
02007-1	37.39	37.55	8.75	8.88	17.04	18.49	2.03	1.85	61.81	63.91	48.22	46.47	0.0147	0.0149	1.23
02007-3	37.39	37.60	8.75	8.64	17.04	18.92	2.03	1.45	61.81	63.36	48.22	46.19	0.0147	0.0150	1.98
02011-3	38.13	38.20	8.39	8.19	39.96	40.63	6.79	6.93	65.78	68.22	48.35	49.38	0.0133	0.0134	0.95
SEP <sup>d</sup>	0.76		0.34		3.70		0.53		1.67		2.09		0.0008		
RPD <sup>e</sup>	2.57		5.01		4.02		7.98		4.83		1.73		6.79		

<sup>a</sup>Experimental value (% g/g). <sup>b</sup>Predictive value (% g/g). <sup>c</sup>Relative deviation (%). <sup>d</sup>Standard error of prediction. <sup>e</sup>Ratio of the standard deviation of the validation set to the standard error of prediction.

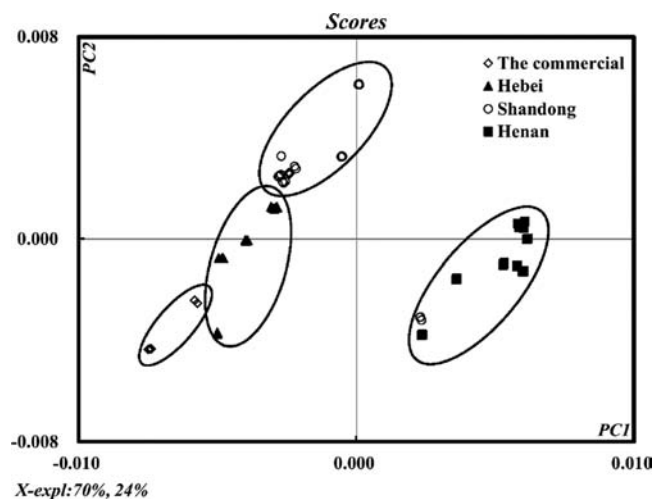


Figure 3. PCA project plot on the NIR of corn steep liquor: samples from different companies were clustered, and the homogeneity of the samples from the same manufacturer was better than that of the samples from different manufacturers. Operated by the Unscrambler (version 9.7 2007).

good correlation coefficients and low predictive errors. The inadequacy of the method is that a third, independent, set from new manufacturers should be predicted to show the power of the best models in this study. However, a new, rapid, and accurate method for determining the main categories of CSL components is established. The method can be used to determine the stability of raw materials and can also be applied in other related fields.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone/fax: +86-10-62781688. E-mail: luoga@mail.tsinghua.edu.cn.

### Author Contributions

<sup>||</sup>Both authors contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We gratefully acknowledge Mr. Zhu and Mr. Tang of the Beijing Kaiyuan Shengshi Science and Technology Development Co., Ltd., for their assistance in the NIR spectroscopic measurements.

## ABBREVIATIONS USED

CD, coefficient of determination; CSL, corn steep liquor; DC, dimension convolution; DM, dry matter; EV, experiment values; HPLC, high-performance liquid chromatography; LV, latent variable; MF, main factors; MSC, multiplicative scatter correction; NIR, near-infrared spectroscopy; PAT, process analytical technology; PCA, principal component analysis; PLS, partial least-squares; PLSR, partial least-squares regression; PV, predictive values;  $R^2$ , correlation coefficient; RD, relative deviation; RMSECV, root-mean-square error of cross-validation; RPD, ratio of standard deviation of the validation set to standard error of prediction; SD, standard deviation; SEC, standard error of calibration; SECV, square error of cross-validation; SEP, standard error of prediction; SPNIRS, spectra pretreatment in near-infrared spectroscopy; TA, total acidity;

TFAA, total free amino acids; TN, total nitrogen; TRS, total reducing sugars; TS, total sugars; TSu, total sulfite.

## REFERENCES

- (1) Juodeikiene, G.; Salomskiene, J.; Eidukonyte, D.; Vidmantiene, D.; Narbutaite, V.; Vaiciulyte, F. L. The impact of novel fermented products containing extruded wheat material on the quality of wheat bread. *Food Technol. Biotechnol.* **2011**, *49* (4), 502–510.
- (2) Watson, A. M.; Gare, M. C.; Kettlewell, P. S.; Brosnan, J. M.; Agu, R. C. Relationships between disease control, green leaf duration, grain quality and the production of alcohol from winter wheat. *J. Sci. Food Agric.* **2010**, *90*, 2602–2607.
- (3) Mona, E. M. M.; Amani, M. D. E. A. Production of  $\beta$ -mannanase by *Bacillus amyloliquefaciens* 10A1 cultured on potato peels. *Afr. J. Biotechnol.* **2008**, *7*, 1123–1128.
- (4) Vendruscolo, F.; Albuquerque, P. M.; Streit, F.; Esposito, E.; Ninow, J. L. Apple pomace: a versatile substrate for biotechnological applications. *Crit. Rev. Biotechnol.* **2008**, *28*, 1–12.
- (5) Claas, H.; Daniel, L.; Thomas, S.; Martin, R. Online-Infrarotspektroskopie in der Bioprozess analytik. *Chem. Ing. Tech.* **2009**, *81* (9), 1385–1396.
- (6) Rathore, A. S.; Bhambure, R.; Ghare, V. Process analytical technology (PAT) for biopharmaceutical products. *Anal. Bioanal. Chem.* **2010**, *398*, 137–154.
- (7) Guan, D. Maize deep processing industries face new challenges. *Fermentation Ind.* **2011**, *229*, 12–14.
- (8) Liggett, R. W.; Koffler, H. Corn steep liquor in microbiology. *Bacteriol. Rev.* **1948**, *12*, 297–311.
- (9) Hull, S. R.; Yang, B. Y.; Venzke, D.; Kulhavy, K.; Montgomery, R. Composition of corn steep water during steeping. *J. Agric. Food Chem.* **1996**, *44* (7), 1857–1863.
- (10) Chovatiya, S. G.; Bhatt, S. S.; Shah, A. R. Evaluation of corn steep liquor as a supplementary feed for *Labeo rohita* (Ham.) fingerlings. *Aquacult. Int.* **2011**, *19* (1), 1–12.
- (11) Nascimento, R. P.; Junior, N. A.; Pereira, N., Jr.; Bon, E. P. S.; Coelho, R. R. Brewer's spent grain and corn steep liquor as substrates for cellulolytic enzymes production by *Streptomyces malaysiensis*. *Lett. Appl. Microbiol.* **2009**, *48*, 529–535.
- (12) Chiani, M.; Akbarzadeh, A.; Farhangi, A.; Mehrabi, M. R. Production of desferrioxamine B (Desferal) using corn steep liquor in *Streptomyces pilosus*. *Pakistan J. Biol. Sci.* **2010**, *13*, 1151–1155.
- (13) Gao, Y.; Yuan, Y. J. Comprehensive quality evaluation of corn steep liquor in 2-keto-L-gulonic acid fermentation. *J. Agric. Food Chem.* **2011**, *59*, 9845–9853.
- (14) Edwinoliver, N. G.; Thirunavukarasu, K.; Purushothaman, S.; Rose, C.; Gowthaman, M. K.; Kamini, N. R. Corn steep liquor as a nutrition adjunct for the production of *Aspergillus niger* lipase and hydrolysis of oils thereof. *J. Agric. Food Chem.* **2009**, *57*, 10658–10663.
- (15) Gentina, J. C.; Acevedo, F.; Villagra, M. P. Short communication: effect of complex nitrogen sources on the production of penicillin acylase by *Bacillus megaterium*. *World J. Microbiol. Biot.* **1997**, *13* (1), 127–128.
- (16) Lu, J. M.; Peng, C.; Ji, X. J.; You, J. Y.; Cong, L. L.; Ouyang, P. K.; Huang, H. Fermentation characteristics of *Mortierella alpine* in response to different nitrogen source. *Appl. Biochem. Biotechnol.* **2011**, *164*, 979–990.
- (17) Badal, C.; Saha, F.; Michael, R. Effects of pH and corn steep liquor variability on mannitol production by *Lactobacillus intermedius* NRRL B-3693. *Appl. Biochem. Biotechnol.* **2010**, *87*, 553–560.
- (18) Bureau, S.; Ruiz, D.; Reich, M.; Gouble, B.; Bertrand, D.; Audergon, J.; Renard, C. G. C. Application of ATR-FTIR for a rapid and simultaneous determination of sugars and organic acids in apricot fruit. *Food Chem.* **2009**, *115*, 1133–1140.
- (19) Louw, E. D.; Theron, K. I. Robust prediction models for quality parameters in Japanese plums (*Prunus salicina* L.) using NIR spectroscopy. *Postharvest Biol. Technol.* **2010**, *58* (3), 176–184.
- (20) Guo, Z. M.; Chen, Q. S.; Chen, L. P.; Huang, W. Q.; Zhang, C.; Zhao, C. J. Optimization of informative spectral variables for the quantification of EGGG in green tea using fourier transform near-infrared (FT-NIR) spectroscopy and multivariate calibration. *Appl. Spectrosc.* **2011**, *65*, 1062–1067.
- (21) Marietta, F.; Agnes, W.; Sandor, T.; Tamas, S. Development of a rapid, non-destructive method for egg content determination in dry pasta using FT-NIR technique. *J. Food Eng.* **2011**, *107*, 195–199.
- (22) Balabin, R.; Smirnov, S. Melamine detection by mid- and near-infrared (MIR/NIR) spectroscopy: a quick and sensitive method for dairy products analysis including liquid milk, infant formula, and milk powder. *Talanta* **2011**, *85*, 562–568.
- (23) Collell, C.; Gou, P.; Arnau, J.; Comaposada, J. Non-destructive estimation of moisture, water activity and NaCl at ham surface during resting and drying using NIR spectroscopy. *Food Chem.* **2011**, *29*, 601–607.
- (24) Fan, L.; Yan, J. Q.; Zhen, C. X.; Liang, Y. L. Research on applying 3,5-dinitrosalicylic acid method to analysing the content of cotton reducing sugar. *Acta Gossypii Sinica* **1996**, *8* (3), 151–154.
- (25) Huang, H.; Hou, Y. Y.; Xiao, X.; Fu, J. Z.; Song, S. Y.; Bai, G.; Luo, G. A. Determination of amino acids in Xuebijing Injection by o-phthalaldehyde-9-fluorenylmethyl chloroformate pre-column derivatization HPLC. *Drugs Clin.* **2011**, *26*, 58–62.
- (26) Lv, L. A.; Shi, T.; Yang, H. H.; Liang, Q. L.; Wang, Y. M.; Luo, G. A. Application of on-line detection and analytical technique by NIR in quality control of water extracting process of Anshen Bunao Liquid. *Chin. Tradit. Herb. Drugs* **2009**, *40*, 224–228.
- (27) Qin, F.; Yang, H. H.; Lv, L. A.; Shi, T.; Liang, Q. L.; Wang, Y. M.; Luo, G. A. Application of near infrared spectra coupled with LLE-PLS modeling to extraction process of Anshen Bunao Syrup. *Chin. Tradit. Patent Med.* **2008**, *30* (10), 1465–1468.
- (28) Yang, H. H.; Qin, F.; Shen, Y. X.; Song, H. K.; Luo, G. A. klsomap – PLS nonlinear modeling method and its application in NIR spectral analysis. *Spectrosc. Spectral Anal.* **2008**, *28*, 149–150.
- (29) Yang, H. H.; Qin, F.; Wang, Y.; Wu, Y. M.; Shi, X. H.; Liang, Q. L.; Wang, Y. M.; Luo, G. A. LLE-PLS nonlinear modeling method for near infrared spectroscopy and its application. *Spectrosc. Spectral Anal.* **2007**, *27*, 1955–1958.
- (30) Mary, D. Q.; Gilvan, W.; Gilcelia, A. C.; Patricio, G. P.; Noemi, N. Determination of simple sugars, malic acid and total phenolic compounds in apple pomace by infrared spectroscopy and PLSR. *Int. J. Food Sci. Technol.* **2010**, *45*, 602–609.
- (31) Wu, W.; Massart, D. L.; Jong, S. D. Kernel-PCA algorithms for wide data part II: Fast cross-validation and application in classification of NIR data. *Chemom. Intell. Lab.* **1997**, *37*, 271–280.
- (32) Jaillais, B.; Pinto, R.; Barros, A. S.; Rutledge, D. N. Outer-product analysis (OPA) using PCA to study the influence of temperature on NIR spectra of water. *Vib. Spectrosc.* **2005**, *39*, 50–58.
- (33) Liu, Y. D.; Sun, X. D.; Ouyang, A. G. Nonstrctive measurement of soluble content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN. *LWT – Food Sci. Technol.* **2010**, *43*, 602–607.
- (34) Zhang, G. Z.; Li, J. S.; Yang, H. T.; Wang, J. N.; Hu, J. D. Determination of amino acid composition in hydrolyzed cotton seed protein and corn steep liquor. *Jiangsu J. Agric. Sci.* **2008**, *24*, 975–976.
- (35) Wang, M.; Zhang, J. S.; Zhao, Y. L.; Zhang, B.; Wang, Y. S.; Guo, H. Z.; Su, Z. G. Determination of free amino acids content in corn syrup by HPLC. *China Brew.* **2010**, *218*, 156–159.
- (36) Chu, X. L.; Xu, Y. P.; Tian, G. Y. Compound with hydroxyl. In *Practical Guide to Interpretive Near-infrared Spectroscopy*; 1st ed.; Chemical Industry Press: Beijing, China, 2009.
- (37) Ozaki, Y.; Liu, Y. L.; Czarniecki, M. A.; Noda, I. FT-NIR spectroscopy of some long-chain fatty acids and alcohols. *Macromol. Symp.* **1995**, *94*, 51–59.