SCIENTIFIC REPORTS

OPEN

Geographical origin traceability of Cabernet Sauvignon wines based on Infrared fingerprint technology combined with chemometrics

Received: 4 March 2019
Accepted: 14 May 2019
Published online: 04 June 2019

Xiao-Zhen Hu[1], Si-Qi Liu[1,2], Xiao-Hong Li[2], Chuan-Xian Wang[2], Xin-Lu Ni[2], Xia Liu[2], Yang Wang[3], Yuan Liu[4] & Chang-Hua Xu[1,5,6,7,8]

Mid-infrared (MIR) and near-infrared (NIR) spectroscopy combined with chemometrics were explored to classify Cabernet Sauvignon wines from different countries (Australia, Chile and China). Commercial wines (n = 540) were scanned in transmission mode using MIR and NIR, and their characteristic fingerprint bands were extracted at 1750-1000 cm^{-1} and 4555-4353 cm^{-1}. Through the identification system of Tri-step infrared spectroscopy, the correlation between macroscopic chemical fingerprints and geographical regions was explored more deeply. Furthermore, Principal component analysis (PCA), soft independent modelling of class analogy (SIMCA) and discriminant analysis (DA) based on MIR and NIR spectra were used to visualize or discriminate differences between samples and to realize geographical origin traceability of Cabernet Sauvignon wines. Through "external test set (n = 157)" validation, SIMCA models correctly classified 97%, 97% and 92% of Australian, Chilean and Chinese Cabernet Sauvignon wines, while the DA models correctly classified 86%, 85% and 77%, respectively. Based on unique digital fingerprints of spectroscopy (FT-MIR and FT-NIR) associated with chemometrics, geographical origin traceability was achieved in a more comprehensive, effective and rapid manner. The developed database models based on IR fingerprint spectroscopy with chemometrics could provide scientific basis and reference for geographical origin traceability of Cabernet Sauvignon wines (Australia, Chile and China).

Geographical origin traceability is of great importance in the process of food quality control. At present, the concept of original geographical indication has become one of the important factors affecting consumers' purchasing preferences in the food and beverage industry[1]. Especially for wine, geographical origin has long been regarded as an inherent criterion in wine identification or classification systems[2,3]. The implementation of comprehensive technologies for origin traceability is urgently needed for consumers, manufacturers, retailers and administrations[1,4,5]. To date, well-established standard means, sensory evaluation, chromatographic and spectrometric methods with expensive instrumentation or time-consuming operation such as gas chromatography (GC), liquid chromatography (LC), mass spectrometry (MS), high performance liquid chromatography (HPLC), elemental and isotopic analysis were applied to discriminate wines on the basis of different geographical origin[1,6–8].

In contrast, molecular spectroscopy technologies have become attractive technique with a bright prospect in applications of batch detection for wine industry[5,9–11], due to environmentally friendly characteristics, as well as the potential savings in analysis time and cost. Infrared spectroscopy that could reflect comprehensive

[1]College of Food Science & Technology, Shanghai Ocean University, Shanghai, 201306, P.R. China. [2]Shanghai Entry-Exit Inspection and Quarantine Bureau, Shanghai, 200135, P.R. China. [3]First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, 300193, P.R. China. [4]School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai, 200240, China. [5]Department of Pharmacology, Yale University, New Haven, CT, 06520, USA. [6]Shanghai Engineering Research Center of Aquatic-Product Processing & Preservation, Shanghai, 201306, P.R. China. [7]Laboratory of Quality and Safety Risk Assessment for Aquatic Products on Storage and Preservation (Shanghai), Ministry of Agriculture, Shanghai, 201306, China. [8]National R&D Branch Center for Freshwater Aquatic Products Processing Technology (Shanghai), Shanghai, 201306, China. Xiaozhen Hu and Siqi Liu contributed equally. Correspondence and requests for materials should be addressed to Y.L. (email: y_liu@sjtu.edu.cn) or C.-H.X. (email: chxu@shou.edu.cn)

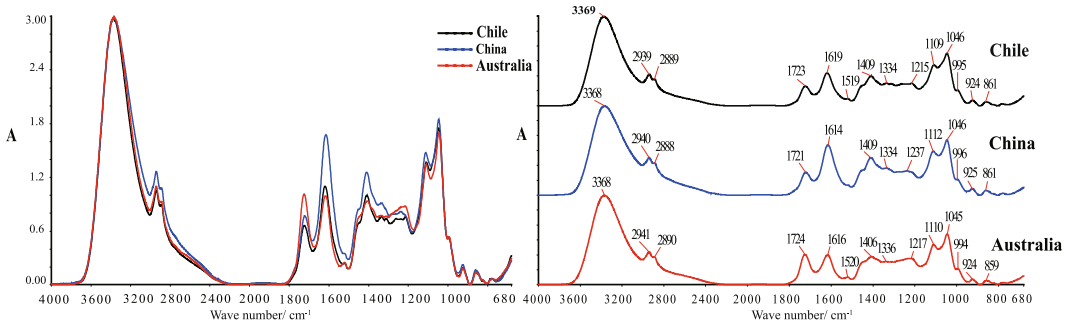

Figure 1. MIR spectra of Cabernet Sauvignon wines from different countries (Chile, China and Australia) in the range of 4000~680 cm^{-1}.

information and behaves as a "fingerprint" of the sample, in the MIR spectral region (4000-400 cm^{-1}), which is caused by the fundamental stretching, bending, and rotating vibrations of the sample molecules, while NIR spectra (12,800-4000 cm^{-1}) results from complex overtones and high-frequency combinations at shorter wavelengths[12–14]. In particular, Tri-step infrared spectroscopy, a comprehensive spectral technique integrating Fourier transform infrared spectroscopy (FT-IR), second derivative infrared spectroscopy (SD-IR) and two-dimensional correlation infrared spectroscopy (2DCOS-IR), has been proved to be an effective technique to reveal main constituents in complicated mixture systems and distinguishing the types and contents of chemical components in highly similar matrices[15–19]. However, few studies focused on the exploration of the key spectral information of origin traceability by recognition mechanism of Tri-step infrared spectroscopy.

Furthermore, multivariate data analysis techniques such as principal component analysis (PCA), cluster analysis (e.g. soft independent modelling of class analogy, SIMCA) and discriminant analysis (DA) can effectively realize the detection of feature patterns or "fingerprint" information related to the geographical origin of samples[5,20], especially for large-scale sample sets. PCA simplifies the data structure by reducing dimension, which is usually used to detect outliers and to identify patterns in the sample distribution before establishment of classification models[21,22]. While SIMCA and DA are supervised classification methods[5,23], which have been commonly used in conjunction with MIR and NIR spectroscopy. Previous studies have demonstrated that both the combined application of MIR-SIMCA and the combination of NIR-DA could provide high predictive ability in the analysis of geographical traceability of wines[9,20,24–26]. Therefore, the application of infrared spectroscopy combined with appropriate chemometrics has potential to realize discrimination and traceability of wines with different geographical origins in a more rapid direct and comprehensive manner.

Cabernet Sauvignon (Vitis vinifera L.) is considered as an ancient and traditional red wine grape variety derived its fame from the south west of France[27]. In emerging grape growing regions called New World wine production countries such as Australia, Chile and China, this variety has become an important red cultivar owing to unique flavor characteristics and broad planting area. Particularly, in China, Cabernet Sauvignon has been currently the most famous red grape variety accepted and favored by wine producers and consumers[28].

Herein, the aim of this study was to explore the macroscopic chemical fingerprints and key spectral information of geographical regions by Tri-step infrared spectroscopy, and further to establish high-throughput classification database models of Cabernet Sauvignon wines with different geographical countries (Australia, Chile, and China) based on unique digital fingerprints of spectroscopy (FT-MIR and FT-NIR) associated with chemometrics.

Results and Discussion

Chemical analysis. Supplementary Table S1 showed the one-way variance analysis (ANOVA) for the chemical constituents of Cabernet Sauvignon wines from different countries (Australia, Chile and China). No statistically significant differences between wine samples analyzed with different countries were observed for Alcohol Content (AC), Glucose plus Fructose (G + F), and Total Phenols (TP). However, the differences in pH, Titratable Acidity (TA) and Volatile Acidity (VA) were statistically significant ($p < 0.05$). By comparing the composition of wines from three countries, Australian Cabernet Sauvignon wines present the highest TA, G + F and TP and the lowest alcohol content, Chilean Cabernet Sauvignon wines present the lowest content of TA, and Chinese Cabernet Sauvignon wines had the highest values of alcohol content, VA and pH, and the lowest content of G + F.

Tri-step IR spectral analysis. *IR spectra of three cabernet sauvignon wines.* According to the one-dimensional FT-MIR spectra (Fig. 1) and the information of corresponding characteristic absorption peaks[29–33] (see Supplementary Table S2), main constituents of Cabernet Sauvignon wine samples from three countries were considered to be the same. Through the overlapping and separated MIR spectra (Fig. 1), significant differences of multiple characteristic peak intensity and shape among wines from three countries (Chile, China and Australia) were observed, such as peaks at 2940, 2890, 1723, 1618, 1409, 1109 and 1046 cm^{-1}, etc. The peak height ratio (1723/1618) of the two main absorption peaks at 1723 and 1618 cm^{-1} were 0.588, 0.475 and 1.091 for Chile, China and Australia (Fig. 1b), which indicated that relative content of the corresponding esters and carboxylic acids were different in wine samples from different countries. Australian Cabernet Sauvignon wines had the largest intensity of ν(C=O) absorption peak at 1723 cm^{-1}, while the peak intensities of stretching vibrational

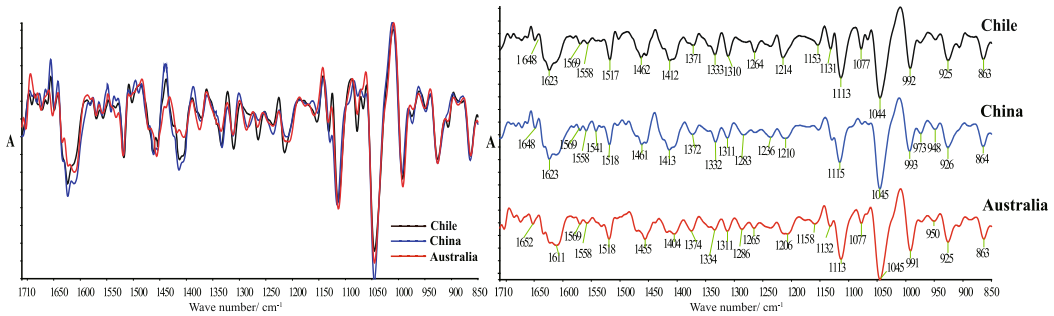

Figure 2. Second derivative spectra of Cabernet Sauvignon wines from different countries (Chile, China and Australia) in the range of 1710~850 cm^{-1}.

absorption peak of COO− at 1618 and 1409 cm^{-1} were significantly weaker than other two countries. Chinese Cabernet Sauvignon wines had the strongest stretching vibration absorption peaks of C−H bond at 2940 and 2890 cm^{-1}, COO− bond at 1618 and 1409 cm^{-1}, and C−O bond at 1107 cm^{-1} and 1046 cm^{-1}, which indicated that the content of alcohol, glycerol and carboxylic acids were higher than the others. Absorption bands around 1275-1200 cm^{-1} were mainly associated with aromatic compounds and their derivatives, ether-containing compounds. In general, according to the absorption peaks intensities (Fig. 1), Australian Cabernet Sauvignon wines present the highest content of ester and carboxylic acids, and the aromatic substances are abundant. Chinese Cabernet Sauvignon wines had less esters, and the carboxylic acid content was much higher than that of the ester. For Chilean Cabernet Sauvignon wines, the esters content was the lowest, while the alcohol and carboxylic acid content were moderate. It has been indicated that Cabernet Sauvignon wines from different countries present unique flavor personality with different flavor components such as alcohol, ester, acids and carbohydrate.

Second derivative IR spectra of three Cabernet Sauvignon wines. Generally, the overlapping absorption peaks can be separated, as well as components with low content or weak absorption intensities in the mixture can be more intuitively identified and compared by SD-IR spectra. Characteristic peaks around 1080-1044 cm^{-1} representing the vibration of C-OH bonds of ethanol, glycerol and sugars (G + F) were observed more directly in the SD-IR spectra of Cabernet Sauvignon wines in the range of 1710-850 cm^{-1} (Fig. 2). Moreover, more information about absorbance peaks appeared, such as peak related to aromatic groups at 1265 cm^{-1}, ν(OC=O), ν(C=C), ν(C−H_2), ν(C−H_3) absorbance peaks presenting organic acids and aldehyde at 1464-1400 cm^{-1}, ν(C=O) peak associated with free amino acids and peptides at 1650 cm^{-1}, and absorption bands of amino acids and their derivatives at 1600-1530 cm^{-1}. According to the macroscopic fingerprint differences (peak intensity, position and shape) in SD-IR spectra of three Cabernet Sauvignon wines, it showed that the amino acid and aromatic compounds including their derivatives (phenols) types and contents of Cabernet Sauvignon wines in three countries are significantly different. Compared with Australia and Chile, sugar and phenols contents of the Cabernet Sauvignon wine samples from China were less.

2DCOS-IR spectra of three cabernet Sauvignon wines. To identify differences among the wines from different countries (Chile, China and Australia) more remarkably and convincingly, the synchronous 2DCOS-IR spectra has been applied in the wave number range of 1800~850 cm^{-1}. In synchronous 2DCOS-IR spectrum, the peaks show the coincidence of the spectral intensity variations at corresponding variables along the perturbation and can be used to verify differences between samples[15]. The auto-peaks on the diagonal line represented the susceptibility and auto-correlativity of certain absorption bands, which produced changes in spectral intensity by thermal treatment. Positive correlation (red/green area) in 2DCOS-IR spectra indicates that a group of absorption bands change simultaneously (either stronger or weaker), while negative correlation (blue area) is completely the opposite[34–36].

Cabernet Sauvignon wines from three countries (Chile, China and Australia) responded differently to thermal perturbation, which made the mean 2DCOS-IR spectra display obvious variation (Fig. 3). The information about positions, relative intensities and correlation of auto-peaks in the temperature ranges of 65–110 °C and 110–120 °C were summarized (see Supplementary Tables S3 and S4). During the thermal perturbation at 65–110 °C, the presence of the strongest auto-peaks at 1648 cm^{-1} and two weak peaks at 1516 cm^{-1} and 1044 cm^{-1} for Chilean Cabernet Sauvignon wines. Australian Cabernet Sauvignon wines had the strongest auto-peak at 1042 cm^{-1} and four weak auto-peaks (1448 cm^{-1}, 1408 cm^{-1}, 1164 cm^{-1}, 965 cm^{-1}), whereas Chinese Cabernet Sauvignon wines had one strong auto-peak at 1044 cm^{-1} and four weak auto-peaks (1721 cm^{-1}, 1560 cm^{-1}, 1512 cm^{-1}, 1164 cm^{-1}). Furthermore, during the thermal perturbation at 110–120 °C, the presence of the strongest auto-peaks at 1580 cm^{-1} and four weak auto-peaks (1435 cm^{-1}, 1250 cm^{-1}, 962 cm^{-1}, 884 cm^{-1}) for Chilean Cabernet Sauvignon wines. Australian Cabernet Sauvignon wines had the strongest auto-peaks at 1160 cm^{-1} and three weak auto-peaks (1408 cm^{-1}, 1152 cm^{-1}, 1042 cm^{-1}), whereas Chinese Cabernet Sauvignon wines had one strong auto-peak at 1650cm^{-1} and three weak auto-peaks (1448 cm^{-1}, 962 cm^{-1}, 884 cm^{-1}). Therefore, Cabernet Sauvignon wines had respective unique fingerprints in the range of 1710-850 cm^{-1} in the synchronous 2DCOS-IR spectra.

Form the all above, the strong automatic peak of Chinese Cabernet Sauvignon wines was mainly the contribution of ethanol and glycerin by the thermal perturbation of 65–110 °C, and the presence of the strongest automatic peak during 110–120 °C was the response of free amino acids and polypeptide components. During the whole

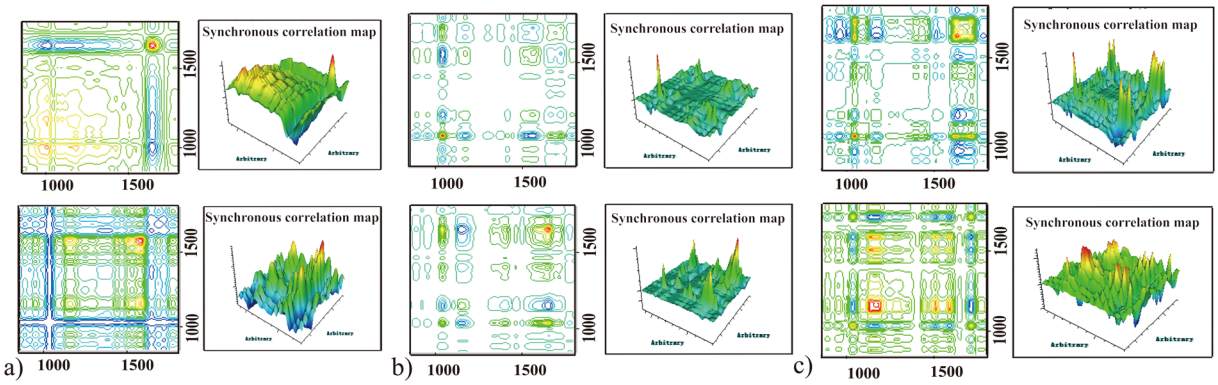

Figure 3. 2DCOS-IR synchronous spectra of Cabernet Sauvignon wines from different countries (Chile, China and Australia) in the range of 1800~850 cm^{-1}.

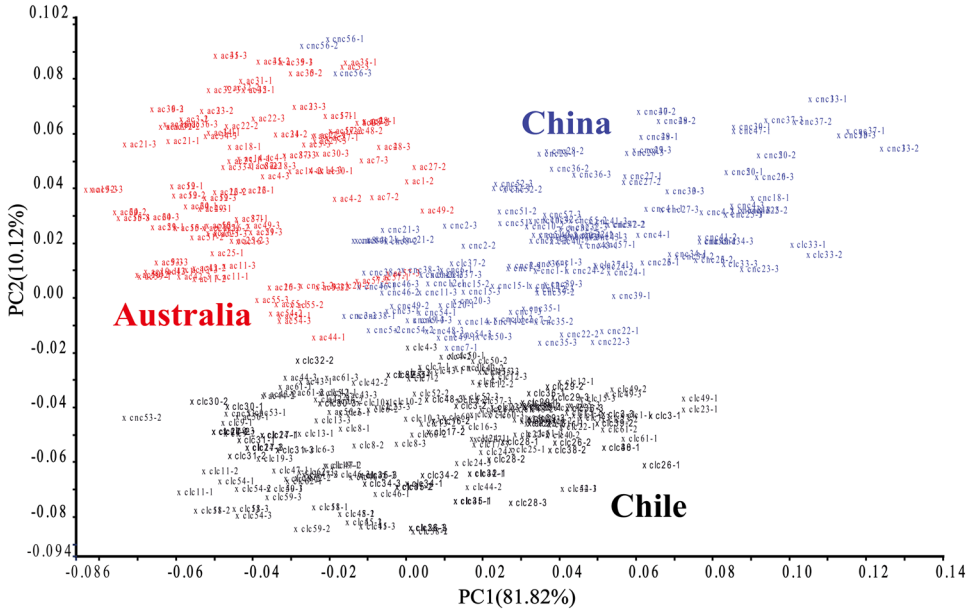

Figure 4. Score plot of the first two PCs for Cabernet Sauvignon wines from different countries (CLC: Chile; CNC: China; AC: Australia) based on FT-MIR spectra.

process of thermal perturbation, the strongest change in temperature response of Chilean wines was primarily amino acid, followed by aromatic substances (phenols) and glycerol components, while Australian wines was mainly contributed by the ethanol, glycerol and phenolic compounds, followed by esters, carboxylic acids and aromatic amino acids.

According to the response intensity of different components to the thermal perturbation in the 2DCOS-IR spectra, geographical differences between the Cabernet Sauvignon wines could be directly judged. Therefore, Chillan, Chinese and Australian wines could be identified and distinguished more clearly and completely based upon the MIR macro-fingerprint characteristics.

PCA analysis. Combining the spectral characteristics extracted from Tri-step IR analysis, the pre-processed MIR spectra (Standard Normal Variate, SNV and De-trending, 1710-850 cm^{-1}) of the Cabernet Sauvignon wine samples were analyzed by PCA (Fig. 4), which could visualize systematic differences between samples from the three countries. The first two PCs displayed almost 92% of total variance in the selected wine samples from three countries. Three clusters representing all wine samples (Chile, China, and Australia) were observed, however, some samples mainly from China and Australia did overlap. To investigate the potential relationship between characteristics of origin and specific chemical composition, the PCA eigenvectors and the spectral characteristics (see Supplementary Fig. S1) were analyzed. PC1 explained 81.82% of the total variance, and the highest loadings were located in the absorption band around 1080-1045 cm^{-1}, which represented the vibration of C-OH band. Additional absorption bonds were observed at 1650 cm^{-1} and 864 cm^{-1}, associated with C=O and −CH bands, respectively. In PC2 (10.12%), the highest eigenvectors were observed at around 1650-1620 cm^{-1} and

Cabernet Sauvignon wines	% Recognition rate (C)	% Rejection rate(C)	% Recognition rate (V)	% Rejection rate (V)
Australia	100 (123/123)	99 (248/251)	91 (49/54)	100 (103/103)
Chile	100 (131/131)	97 (236/243)	89 (49/55)	100 (102/102)
China	100 (120/120)	100 (254/254)	73 (35/48)	100 (109/109)

Table 1. Classification performance report of Calibration (C) and Validation (V) for discriminating Cabernet Sauvignon wines from different countries based on MIR.

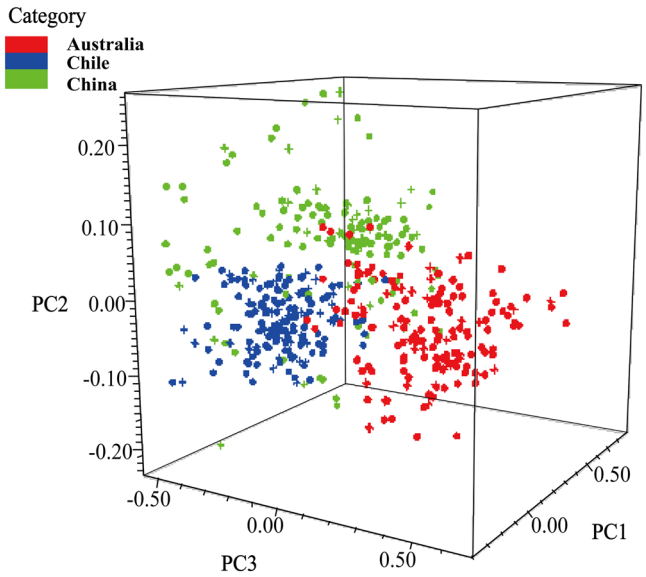

Figure 5. Three-dimensional PCA plot for Cabernet Sauvignon wines from different countries (Australia (AC), Chile (CLC) and China (CNC)) based on NIR spectra.

1100-980 cm^{-1}. Therefore, the PCA analysis for Australian, Chillan and Chinese Cabernet Sauvignon wines were performed based on spectral fingerprint information of MIR. It further demonstrated that alcohols, carbohydrates (glucose and fructose), organic acid and phenolic compounds contributed the strongest differences among the Cabernet Sauvignon wines from different countries (Australia, Chile and China).

Cluster analysis. In order to realize cluster analysis of Cabernet Sauvignon wine samples from three countries, SIMCA was performed on the spectral characteristics of wine samples extracted by PCA. 540 parallel samples were objectively classified, and 157 samples (54 samples for Australia, 55 for Chile, 48 for China) were randomly selected as external validation set. The parameters of SIMCA model were summarized (Supplementary Table S5). The between-class distance between the Cabernet Sauvignon wine samples from three countries was >1 (Supplementary Table S6), which indicated that the degree of separation between various categories was relatively large, and the difference of categories was obvious in the SIMCA model. In addition, the reliability of classification model has been verified by recognition and rejection rate.

According to the classification result of SIMCA model (Table 1, Supplementary Table S9), the recognition rate of calibration set and the rejection rate of validation set in samples with three countries were 100%, which demonstrated that the sensitivity of the calibration set and the specificity of the validation set in SIMCA model were accurate. Nevertheless, the recognition rate of the verification set in Chinese samples was only 73%. Correct classification rates for Australian, Chilean and Chinese wines were 97%, 97%, and 92%, respectively. It has been showed that Cabernet Sauvignon wines from three countries were classified effectively by MIR coupled with PCA-based SIMCA.

Discriminant analysis (DA). Supplementary Tablesnon wine samples (406 samples for the calibration set and 134 samples for the validation set) were analyzed using the Mahalanobis distance discriminant analysis (DA). According to the NIR spectral information and optimization analysis (see Supplementary Tables S7[2,37–39] and S8), the pre-processed NIR spectra (SNV, 4555-4353 cm^{-1}) of the Cabernet Sauvignon wine samples were selected for modeling (Supplementary Fig. S2). Through the projection and distribution of the wine samples in the three-dimensional feature space (Fig. 5), the samples of three countries were distinguished. Furthermore, the clustering trend based on FT-NIR is basically consistent with the results of FT-MIR.

Cabernet Sauvignon wine samples from three geographical countries were classified based on Mahalanobis distance discrimination (Fig. 6). Correct classification rates for Australian, Chilean and Chinese Cabernet Sauvignon wines were 86%, 85% and 77% (Table 2), respectively. Consistent with SIMCA model based on

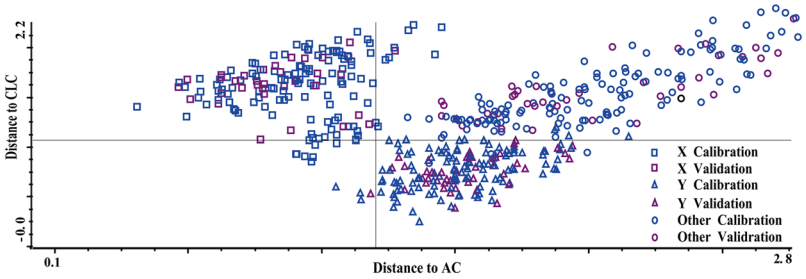

Figure 6. Classification of three Cabernet Sauvignon wines from different countries using Mahalanobis distance discriminant analysis method (□)AC; (△)CLC; (○) CNC.

Cabernet Sauvignon wines	% Recognition rate (C)	% Rejection rate (C)	% Recognition rate(V)	% Rejection rate (V)
Australia	87 (117/135)	100 (258/258)	84 (37/44)	100 (90/90)
Chile	85 (111/130)	91 (240/263)	84 (46/55)	89 (70/79)
China	78 (100/128)	91 (241/265)	71 (25/35)	88 (87/99)

Table 2. Classification performance report of Calibration (C) and Validation (V) for discriminating Cabernet Sauvignon wines from different countries based on NIR.

FT-MIR, prediction effects of Australian and Chilean wines were better than Chinese wines. Furthermore, the accuracy and sensitivity of the SIMCA model based on FT-MIR were better than that of the DA model using FT-NIR.

Conclusion

In this study, we have attempt to establish high-throughput classification models of Cabernet Sauvignon wines with three geographical countries based on unique digital fingerprints of spectroscopy (FT-MIR and FT-NIR) associated with chemometrics.

Through the identification system of Tri-step mid infrared spectroscopy technology, with chemical analysis results of alcohol, pH, total acid, volatile acid, total phenol, glucose and fructose as reference, the macroscopic characteristic fingerprint bands of different countries were extracted at 1750-1000 cm^{-1}. As the increasing resolution of Tri-step infrared spectroscopy, apparent differences in the Cabernet Sauvignon wine samples have been fully visualized due to the fingerprint information (positions and relative intensities of characteristic peaks). Three Cabernet Sauvignon wines from different countries were successfully discriminated and classified in a rapid and holistic manner. Moreover, 540 Cabernet Sauvignon wine samples have been objectively used by chemometrics (PCA, SIMCA and DA) based on MIR and NIR macro-fingerprints to realize rapid traceability analysis of unknown Cabernet Sauvignon wine samples from Australia, Chile and China. These results suggested that the prediction effects of Australian and Chilean wines were better than that of Chinese wines, and the classification effect of SIMCA model based on FT-MIR was more precise than DA model based on FT-NIR. For the detection of Chinese wines, it is necessary to improve the accuracy of the classification by establishing robust models with more samples.

Conclusively, it has been demonstrated that the developed database models based on FT-MIR and FT-NIR coupled with chemometrics (PCA, SIMCA and DA) could be applied as a reference for geographical origin traceability of Cabernet Sauvignon wines (Australia, Chile and China) in a more comprehensive, effective and rapid manner.

Experimental

Samples and reference data. Commercial Cabernet Sauvignon wine samples from Australia (South Australia, Western Australia, Victoria and New South Wales, respectively), Chile (Central Valley) and China (Hebei, Shandong and Ningxia, respectively) were collected from commercial wineries and Shanghai Entry-Exit Inspection and Quarantine Bureau. A total of 540 wine samples includig Australian (61 labels × 3 bottles), Chilean (62 labels × 3 bottles) wines and Chinese (57 labels × 3 bottles) wines were used in this study. All samples were from 2010 to 2016 vintages, produced in the main wineries of the geographical zones: Penfolds, Capel Vale, McWilliam's, Vina Ventisquero, Concha y Toro, Montes, Baron Philippe de Rothschild, ChangYu, Greatwall Manor, Chateau Bacchus, Imperial Horse and Dynasty. Reference data for alcohol content (AC), glucose and fructose (G + F), pH, titratable acidity (TA), volatile acidity (VA) and total phenols (TP) were obtained using standard methods (GB 15037-2006)[31,40].

Instrument. FT-IR spectrometer (Spotlight 400, PerkinElmer, UK) equipped with a deuterated triglycine sulfate (DTGS) detector and Universal ATR sampling accessory. Thermo Scientific Nicolet iS5 FT-IR spectrometer equipped with ATR temperature controller which performed the thermal perturbation was used to obtain the

two-dimensional correlation spectra. The IR spectra were recorded from 4000 to 400 cm^{-1} and 12,800-4000 cm^{-1}. Spectra were recorded with 32 scans and 0.5 cm/s^{-1} of OPD speed.

Other equipments include ultrapure water machine (Milli-Q Reference, Elix Reference), Seven Excellence S400-B multifunction tester (METTLER TOLEDO, Switzerland), LC-20A high performance liquid chromatograph (SHIMADZU, Japan), T70 full-automatic potentiometric titrometer (0.01 pH) (METTLER TOLED, Switzerland), Evolution™ 300 UV-Visible spectrometer (Thermo Fisher Scientific, USA), DMA4500 digital densitometer (Anton-Paar, Austria) and rotary evaporator (IK ARV10 basic, HB10 digital, German) and Freeze Dryer BTP-3XLOVX (Virtis, American).

Procedure. *Spectroscopic measurements of FT-MIR.* 3 mL of each wine sample was taken from freshly opened bottles and distilled at 40 °C for about 8 min by rotating evaporator until the resulting wine sample was essentially alcohol-free. Then the resulting sample was freeze-dried for 24 h, each sample (1~2 mg) was mixed with KBr (100 mg) into powder and finally pressed into tablets. FT-IR spectra of samples were scanned at room temperature by PerkinElmer FT-IR spectrometer in transmission mode. Each spectrum was recorded as the average of 32 scans with 4 cm^{-1} resolution in the wavenumber range of 4000-400 cm^{-1}. The SD-IR spectra was obtained by Savitzky-Golay polynomial fitting (13-point smoothing) with PerkinElmer Spectrum software (Version 10.4.3).

In order to obtain 2DCOS-IR spectra representing the overall difference of selected samples, each wine sample from three countries was placed in ATR accessory connected with the temperature controller and recorded in two variable temperature-gradients: from 65 to 110 °C with an increasing rate at 2 °C/min at an interval of 10 °C, and from 110 to 120 °C with an increasing rate at 2 °C/min at an interval of 2 °C. After removing the abnormal sample information, a series of mean dynamic spectra of three countries in variable temperature-gradients were processed using 2DCOS-IR correlation analysis software (Thermo Scientific, Nicolet iN10 SpectraCorr). Then, the mean 2DCOS-IR spectra of three countries were obtained.

Spectroscopic measurements of FT-NIR. Wine Samples were scanned in transmission mode using near-IR fiber-optic probe accessory of Nicolet iS50 (Thermo Fisher Scientific, America) equipped with an indium-gallium-arsenide (InGaAs) detector. Each spectrum was recorded as the average of 32 scans with 4 cm^{-1} resolution in the wavenumber range of 12,800-4000 cm^{-1}. Spectra of all samples were collected at room temperature with reference background spectrum recorded using air. All the raw FT-NIR data were processed with Omnic spectrum software (Version 9.2.106).

Multivariate data analysis. A total of 540 dry wine samples (183 samples for Australia, 186 for Chile, 171 for China) were used to establish models of geographical origin traceability by PCA, SIMCA and DA. PCA was conducted based on FT-MIR spectra (1710-850 cm^{-1}) by Spectrum Quant+ v 4.6.0.0176 (Perkin Elmer Inc.). SIMCA models (FT-MIR) based on PCA were performed by Assure ID v 4.1.0.0195 (PerkinElmer Inc.). And Mahalanobis distance discriminant models (FT-NIR) based on PCA were developed by TQ Analyst v 9.7.179 (Thermo Fisher Scientific Inc.). Statistical analysis of chemical components was analyzed by SPSS Statistics (Version 23.0) software.

Data Availability

The dataset generated or analyzed during the current study are available from corresponding author upon reasonable request.

References

1. Luykx, D. M. A. M. & Van Ruth, S. M. An overview of analytical methods for determining the geographical origin of food products. *Food chemistry* **107**, 897–911, https://doi.org/10.1016/j.foodchem.2007.09.038 (2008).
2. Cozzolino, D., Smyth, H. E. & Gishen, M. Feasibility study on the use of visible and near-infrared Spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. *Journal of Agricultural and Food Chemistry* **51**, 7703–7708, https://doi.org/10.1021/jf034959s (2003).
3. Green, J. A., Parr, W. V., Breitmeyer, J., Valentin, D. & Sherlock, R. Sensory and chemical characterisation of Sauvignon blanc wine: Influence of source of origin. *Food Research International* **44**, 2788–2797, https://doi.org/10.1016/j.foodres.2011.06.005 (2011).
4. Tregear, A., Kuznesof, S. & Moxey, A. Policy initiatives for regional foods: some insights from consumer research. *Food Policy* **23**, 383–394, https://doi.org/10.1016/s0306-9192(98)00044-x (1998).
5. Cozzolino, D., Cynkar, W. U., Shah, N. & Smith, P. A. Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food chemistry* **126**, 673–678, https://doi.org/10.1016/j.foodchem.2010.11.005 (2011).
6. Kamiloglu, S. Authenticity and traceability in beverages. *Food chemistry* **277**, 12–24, https://doi.org/10.1016/j.foodchem.2018.10.091 (2019).
7. Cordella, C., Moussa, I., Martel, A. C., Sbirrazzuoli, N. & Lizzani-Cuvelier, L. Recent developments in food characterization and adulteration detection: Technique-oriented perspectives. *Journal of Agricultural and Food Chemistry* **50**, 1751–1764, https://doi.org/10.1021/jf011096z (2002).
8. Jiang, B., Xi, Z., Luo, M. & Zhang, Z. Comparison on aroma compounds in Cabernet Sauvignon and Merlot wines from four wine grape-growing regions in China. *Food Research International* **51**, 482–489, https://doi.org/10.1016/j.foodres.2013.01.001 (2013).
9. Liu, L., Cozzolino, D., Cynkar, W. U., Gishen, M. & Colby, C. B. Geographic classification of Spanish and Australian tempranillo red wines by visible and near-infrared spectroscopy combined with multivariate analysis. *Journal of Agricultural and Food Chemistry* **54**, 6754–6759, https://doi.org/10.1021/jf061528b (2006).
10. Mandrile, L., Zeppa, G., Giovannozzi, A. M. & Rossi, A. M. Controlling protected designation of origin of wine by Raman spectroscopy. *Food chemistry* **211**, 260–267, https://doi.org/10.1016/j.foodchem.2016.05.011 (2016).
11. Chandra, S., Chapman, J., Power, A., Roberts, J. & Cozzolino, D. Origin and Regionality of Wines—the Role of Molecular Spectroscopy. *Food Analytical Methods* **10**, 3947–3955, https://doi.org/10.1007/s12161-017-0968-1 (2017).
12. Cozzolino, D., Cynkar, W., Shah, N. & Smith, P. Technical solutions for analysis of grape juice, must, and wine: the role of infrared spectroscopy and chemometrics. *Analytical & Bioanalytical Chemistry* **401**, 1475–1484 (2011).
13. Smyth, H. & Cozzolino, D. Instrumental Methods (Spectroscopy, Electronic Nose, and Tongue) As Tools To Predict Taste and Aroma in Beverages: Advantages and Limitations. *Chemical reviews* **113**, 1429–1440, https://doi.org/10.1021/cr300076c (2013).

14. Hou, S.-W. *et al*. Integrated recognition and quantitative detection of starch in surimi by infrared spectroscopy and spectroscopic imaging. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **215**, 1–8, https://doi.org/10.1016/j.saa.2019.02.080 (2019).
15. Liu, S. *et al*. Rapid identification of pearl powder from Hyriopsis cumingii by Tri-step infrared spectroscopy combined with computer vision technology. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **189**, 265–274, https://doi.org/10.1016/j.saa.2017.08.031 (2018).
16. Xu, C.-H., Liu, S.-L., Zhao, S.-N., Li, S.-Z. & Sun, S.-Q. Unveiling Ontogenesis of Herbal Medicine in Plant Chemical Profiles by Infrared Macro-Fingerprinting. *Planta Medica* **79**, 1068–1076, https://doi.org/10.1055/s-0032-1328764 (2013).
17. Hu, W. *et al*. Rapid Discrimination of Different Grades of White Croaker Surimi by Tri-Step Infrared Spectroscopy Combined with Soft Independent Modeling of Class Analogy (SIMCA). *Food Analytical. Methods* **9**, 831–839, https://doi.org/10.1007/s12161-015-0258-8 (2016).
18. Gu, D. C. *et al*. A rapid analytical and quantitative evaluation of formaldehyde in squid based on Tri-step IR and partial least squares (PLS). *Food chemistry* **229**, 458–463, https://doi.org/10.1016/j.foodchem.2017.02.082 (2017).
19. Zhu, L. *et al*. (2018) Rapid Quality Discrimination and Amino Nitrogen Quantitative Evaluation of Soy Sauces by Tri-Step IR and E-nose. *Food Analytical Methods* **11**(11), 3201–3210, https://doi.org/10.1007/s12161-018-1284-0 (2018).
20. Liu, L. *et al*. Preliminary study on the application of visible-near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. *Food chemistry* **106**, 781–786, https://doi.org/10.1016/j.foodchem.2007.06.015 (2008).
21. Ríos-Reina, R., García-González, D. L., Callejón, R. M. & Amigo, J. M. NIR spectroscopy and chemometrics for the typification of Spanish wine vinegars with a protected designation of origin. *Food Control* **89**, 108–116, https://doi.org/10.1016/j.foodcont.2018.01.031 (2018).
22. Ziegel, E. R. A User-Friendly Guide to Multivariate Calibration and Classification. *Technometrics* **46**, 3 (2004).
23. Gu, D.-C. *et al*. A novel method for rapid quantitative evaluating formaldehyde in squid based on electronic nose. *LWT* **101**, 382–388, https://doi.org/10.1016/j.lwt.2018.11.012 (2019).
24. Urickova, V. & Sadecka, J. Determination of geographical origin of alcoholic beverages using ultraviolet, visible and infrared spectroscopy: A review. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy* **148**, 131–137, https://doi.org/10.1016/j.saa.2015.03.111 (2015).
25. Urbano Cuadrado, M., Luque de Castro, M. D. & Gomez-Nieto, M. A. Study of spectral analytical data using fingerprints and scaled similarity measurements. *Analytical and bioanalytical chemistry* **381**, 953–963, https://doi.org/10.1007/s00216-004-2954-x (2005).
26. Shen, F. *et al*. Discrimination Between Shaoxing Wines and Other Chinese Rice Wines by Near-Infrared Spectroscopy and Chemometrics. *Food and Bioprocess Technology* **5**, 786–795, https://doi.org/10.1007/s11947-010-0347-z (2010).
27. Bowers, J. E. & Meredith, C. P. The parentage of a classic wine grape, Cabernet Sauvignon. *Nature Genetics* **16**, 84–87, https://doi.org/10.1038/ng0597-84 (1997).
28. Jiang, B. & Zhang, Z.-W. Comparison on Phenolic Compounds and Antioxidant Properties of Cabernet Sauvignon and Merlot Wines from Four Wine Grape-Growing Regions in China. *Molecules* **17**, 8804–8821, https://doi.org/10.3390/molecules17088804 (2012).
29. Cozzolino, D., Holdstock, M., Dambergs, R. G., Cynkar, W. U. & Smith, P. A. Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia. *Food chemistry* **116**, 761–765, https://doi.org/10.1016/j.foodchem.2009.03.022 (2009).
30. Cozzolino, D., Cynkar, W., Shah, N. & Smith, P. Feasibility study on the use of attenuated total reflectance mid-infrared for analysis of compositional parameters in wine. *Food Research International* **44**, 181–186, https://doi.org/10.1016/j.foodres.2010.10.043 (2011).
31. Zhang, Y.-l *et al*. Discrimination of different red wine by Fourier-transform infrared and two-dimensional infrared correlation spectroscopy. *Journal of Molecular Structure* **974**, 144–150, https://doi.org/10.1016/j.molstruc.2010.03.021 (2010).
32. Cocciardi, R. A., Ismail, A. A. & Sedman, J. Investigation of the potential utility of single-bounce attenuated total reflectance Fourier transform infrared spectroscopy in the analysis of distilled liquors and wines. *Journal of Agricultural and Food Chemistry* **53**, 2803–2809, https://doi.org/10.1021/jf048663d (2005).
33. Banc, R., Loghin, F., Miere, D., Fetea, F. & Socaciu, C. Romanian wines quality and authenticity using FT-MIR spectroscopy coupled with multivariate data analysis. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* **42**, 556–564 (2014).
34. Guo, X.-X. *et al*. Rapid analysis and quantification of fluorescent brighteners in wheat flour by Tri-step infrared spectroscopy and computer vision technology. *Journal of Molecular Structure* **1099**, 393–398, https://doi.org/10.1016/j.molstruc.2015.06.081 (2015).
35. Zhang, X. *et al*. Accelerated chemotaxonomic discrimination of marine fish surimi based on Tri-step FT-IR spectroscopy and electronic sensory. *Food Control* **73**, 1124–1133, https://doi.org/10.1016/j.foodcont.2016.10.030 (2017).
36. Gan, J.-H. *et al*. Analysis and discrimination of ten different sponges by multi-step infrared spectroscopy. *Chinese Chemical Letters* **26**, 215–220, https://doi.org/10.1016/j.cclet.2015.01.012 (2015).
37. Cozzolino, D. *et al*. Analysis of elements in wine using near infrared spectroscopy and partial least squares regression. *Talanta* **74**, 711–716, https://doi.org/10.1016/j.talanta.2007.06.045 (2008).
38. Li, S., Wilkinson, K. L. & Cozzolino, D. The use of near infrared reflectance spectroscopy to identify the origin of oak shavings used in wine aging. *Journal of Food Measurement and Characterization* **8**, 356–361, https://doi.org/10.1007/s11694-014-9196-1 (2014).
39. Di Egidio, V., Sinelli, N., Giovanelli, G., Moles, A. & Casiraghi, E. NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. *European Food Research and Technology* **230**, 947–955, https://doi.org/10.1007/s00217-010-1227-5 (2010).
40. Industry, C. I. O. F. F. *et al*. Vol. GB/T 15037-2006 (General administration of quality supervision, inspection and quarantine of the People's Republic of China) (2006).

Acknowledgements

This study was funded by Shanghai Pujiang Program (No. 18PJ1432600), National Key Research and Development Program of China (No. 2016YFD0401501), Special development fund project of China (Shanghai) Pilot Free Trade Zone Administration Shanghai FTZ Authority (SHFTZ No. [2016] 192), the Technology Standard Project of Shanghai Municipal Science and Technology Commission (No. 14DZ0501001), AQSIQ Project (No. 2015IK226). The authors are grateful to spiritual support from Mayday.

Author Contributions

X.Z.H. and S.Q.L. are involved in the experimental operation, statistical analysis and manuscript writing, and their contribution to this work is equal. X.H.L. and C.X.W. assisted in experimental research and revision of manuscripts. X.L.N. and X.L. helped analyze the IR datas. Y.W. developed the analytical method. Y.L. and C.H.X. provided research ideas for this work and helped with the design.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-44521-8.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.