Advance Journal of Food Science and Technology 12(3): 150-154, 2016
DOI:10.19026/ajfst.12.2872
ISSN: 2042-4868; e-ISSN: 2042-4876

Submitted: September 26, 2015 Accepted: October 30, 2015 Published: September 25, 2016

Research Article

Identification of Dark Tea (*Camellia sinensis* (*L.*)) Origins According to Chemical Composition Combined with Bayes Classification Pattern Recognition

Jingming Ning, Junting Fang, Xianjingli Luo, Zhengzhu Zhang and Xiaochun Wan
State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036, P.R. China

Abstract: As one of the six major teas in China, dark tea is mainly produced in Yunnan, Hunan, Sichuan, Hubei and Guangxi provinces of china. At present, identification geographical of teas mainly depends on the sensory evaluation, because of lacking the quantitative discriminate method. In this study, 38 dark teas were taken, which were collected from five regions. And the main chemical compositions of tea samples were detected according to international standard. Using SPSS18.0 statistical software to reduction dimension, then chose four compositions (GA, EGC, caffeine, total catechins) as the principal component factors, by using Bayes discriminate analysis method, we established the quantitative discriminate model, which could identify the dark teas from different regions. The results show that the Bayes discriminate analysis can be used to discriminate the 38 samples from five regions and the correct rate could be reached 100%, which means the methods established is reliable.

Keywords: Bayes, Dark tea, different origins, pattern recognition

INTRODUCTION

Dark tea is one of the six major teas in China, which initial processes are fixing, rolling, piling and drying. As the key step of forming dark tea characteristics, piling is based on the activity of microorganism, promotion of a series of complex biochemical changes through biochemical dynamics (extracellular enzyme), materialized power (microbial heat combined with moisture of green tea) and the combined effect of microbial metabolism itself (Wan, 2003; Zengsheng *et al*., 1991; Fulin *et al*., 2006). Therefore, the common feature of dark tea manufactured from different regions is its old and rough raw material, coarse shape, old stems and pile discoloration. In that case, it is difficult to discriminate by sensory evaluation without the compression molding.

At present, Fourier Transform Infrared Spectroscopy (FTIR) has been used to identify different regions of dark teas. 5 typical dark tea samples were analyzed by Fourier Transforms Infrared Spectroscopy (FTIR) and their spectral characteristics were also compared (Xin-He *et al*., 2012). Infrared absorption spectrometry was used to compare the infrared absorption spectral differences of 37 dark tea samples from 6 different regions. And the characteristic peaks of the FTIR fingerprinting of dark tea samples were recognized and compared (Yan and Gui-nian, 2014). Therefore, different recognition methods of the fingerprint for solar dried Pu'er raw tea were investigated in order to identify Pu'er raw tea from different places (JingMing *et al*., 2010). The electronic tongue (He *et al*., 2009) and micellar electrokinetic capillary Chromatography (Kodama *et al*., 2007) was also used to identify tea of different regions. But these methods above all are qualitative, not give us quantitative discriminant method on the region of tea discrimination.

In this study, take 38 representative dark teas as material, which were collected from five regions. And the main chemical compositions, which were Gallic Acid (GA), (−)-Epicatechin (EC), (+)-Catechin (+C), (−)-Epicatechin Gallate (ECG), (−)-epigallocatechin (EGC), (−) -Epigallocatechin-Gallate(EGCG), caffeine and theanine were detected according to international standard. Using SPSS18.0 statistical software to reduction dimension, then choose four compositions (GA, EGC, caffeine, total catechins) as the principal component factors, by using Bayes discriminant analysis method, we established the quantitative discriminant model which can identify the dark tea from different regions.

Corresponding Author: Xiaochun Wan, State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036, P.R. China

MATERIALS AND METHODS

Tea samples: Thirty eight kinds of representative dark tea samples manufactured in five regions were collected, in which 18 Yunnan Pu'er tea samples, 4 Hunan Fuzhuan tea samples, 6 Guangxi Liubao tea samples, 5 Hubei Qingzhuan tea samples and 5 Sichuan Kangzhuan tea samples.

Test method: The detection method of the main chemical compositions GA, EC, +C, ECG, EGC, EGCG, caffeine and theanine of tea samples were determined by ISO (ISO 14502-2:2005; ISO 10727:2002; ISO/WD19563).

Data processing methods:

Principal component analysis: This technique can be applied to generalize varieties of information in large amount of measurement data, which is a effective method for the dimensionality reduction (Wei-ping *et al*., 2007; Duda, 2003; Shushen and Zhongsheng, 1999). PCA is based on the derivation of linear combinations of the variables to produce new variables called Principle Components (PCs) that are uncorrelated (Wold *et al*., 1987).

Bayes discriminant analysis: Bayes discriminant is a probability discriminant analysis in statistical classification field and begins with a prior distribution over all possible models by using probability statistical knowledge classification:

$$W = b_0+b_1x_1+b_2x_2+\ldots+b_px_p$$

Based on the distribution density functions, by calculating the probability of the X sample from a general in the observation of the conditions of X sample, discriminate which classification the X sample belongs to Wei (2011).

RESULTS AND ANALYSIS

Quantitative analysis of chemical composition: High-Performance Liquid Chromatography (HPLC) was employed to analyze the quantization of the dark tea catechins and caffeine. Gallic acid, caffeine as well as various kinds of catechin in tea samples could be well separated (Fig. 1a). Peak symmetry and well abundance showed this method was a reliable and effective approach for routine analysis of dark tea. The content of +C and EC were too low to detect in some of dark teas.

Different concentrations of caffeine were prepared to be analyzed on HPLC, obtaining the caffeine standard curve based on the peak area. According to GA and catechins relative caffeine correction factor (ISO 14502-2:2005), they were quantitative analyzed, respectively. Caffeine standard curve was as following: Y = 12816x, R^2 = 1.0000. Representative chromatogram of theanine of dark tea was shown in Fig. 1b. Different concentrations of theanine were prepared by ISO,

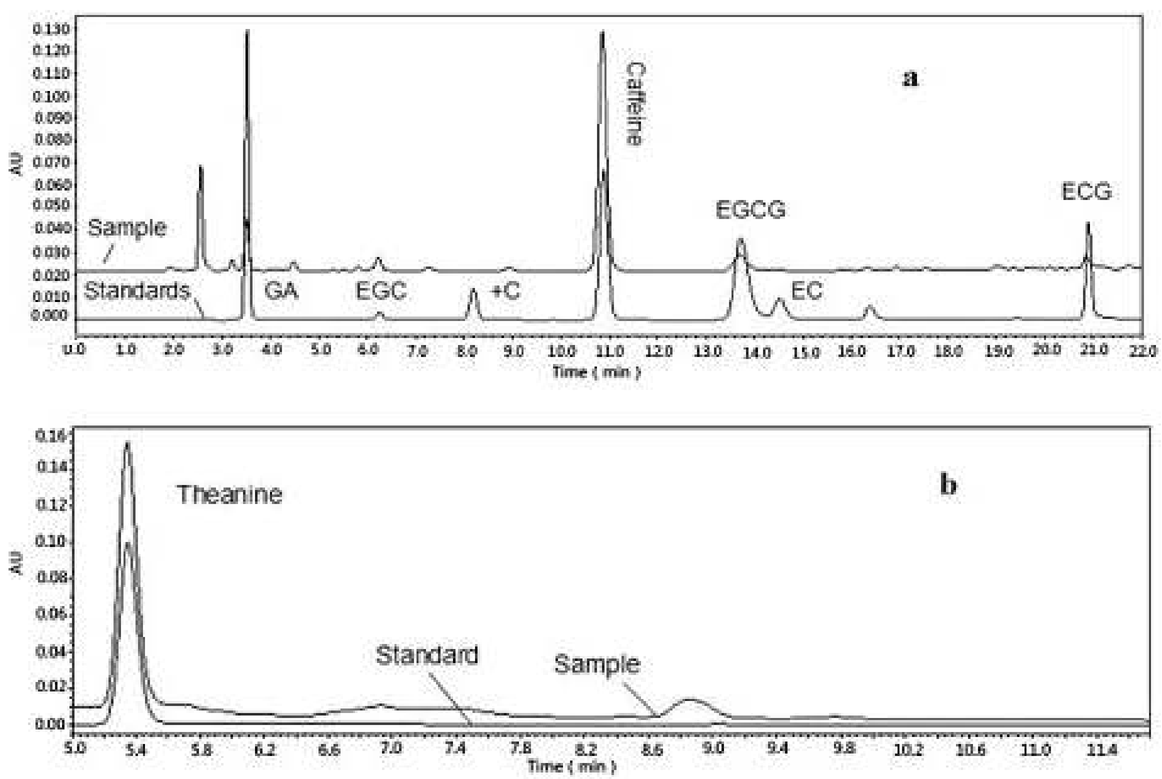

Fig. 1: HPLC chromatogram of mixed references of catechins, caffeine and theanine

Table 1: Contents of main components in different area dark tea

	Origins				
Contents	Yunnan	Hunan	Guangxi	Hubei	Sichuan
GA	0.47±0.28	0.30±0.03	0.19±0.12	0.14±0.05	0.28±0.06
EGC	0.13±0.13	0.85±0.11	0.86±0.53	0.51±0.22	0.15±0.06
+C	0.03±0.03	0.06±0.02	0.05±0.03	0.01±0.01	0.01±0.01
Caffeine	3.62±0.12	2.41±0.28	3.45±0.38	1.25±0.18	1.61±0.20
EGCG	0.01±0.02	1.32±0.30	0.09±0.06	0.44±0.26	0.16±0.16
EC	0.12±0.11	0.50±0.13	0.20±0.11	0.13±0.07	0.03±0.03
ECG	0.03±0.04	0.59±0.13	0.08±0.04	0.15±0.08	0.08±0.04
Total catechins	0.31±0.29	3.32±0.48	1.27±0.71	1.23±0.61	0.43±0.29
Theanine	0.01±0.02	0.14±0.04	0.01±0.01	0.06±0.07	0.05±0.03

Table 2: Contribution rate and cumulative contribution rate of each principal component

	Initial eigen values		
Principal component number	Proportion	Variance%	Cumulative%
1	5.109	56.772	56.772
2	1.989	22.099	78.871
3	0.974	10.817	89.688
4	0.466	5.174	94.862
5	0.210	2.332	97.195
6	0.204	2.269	99.463
7	0.034	0.377	99.841
8	0.014	0.159	100.000
9	2.554E-16	2.837E-15	100.000

obtaining the theanine standard curve by the peak area. Theanine standard curve was as following: $Y = 7851.1x+3089$, $R^2 = 0.9998$.

Results from chemical compositions of dark tea samples: High-Performance Liquid Chromatography (HPLC) was employed to analyze the beneficial components of those dark tea samples. The contents of main components (GA, EC, +C, ECG, EGC, EGCG, caffeine and theanine) were shown in Table 1. In dark tea samples, the amounts of total catechins ranged between 0.31 and 3.32%, while the amounts of caffeine in different samples ranged from 1.25 to 3.62%. The highest content of total catechins was shown in Hunan which was 3.32%. The lowest content of total catechins was shown in Yunnan which was 0.31%. EGCG and EGC were the most abundant catechins in the majority of the dark tea. The levels of EGCG ranged from 0.01% (Yunnan) to 1.32% (Hunan) and the levels of EGC ranged from 0.13% (Yunnan) to 0.86% (Guangxi). The main catechins in Hunan were EGCG and EGC which accounted for about 39.8 and 25.6%. In Hubei dark teas, EGC and EGCG accounted for about 41.5 and 35.8%. Theanine varied from 0.01 to 0.14% and the levels of GA ranged from 0.14% (Hubei) to 0.47% (Yunnan).

Results of principal component analysis: Nine kinds of standardized primitive variables (GA, EGC, +C, Caffeine, EGCG, EC, ECG, total catechins, theanine were treated as X_1, X_2, X_3……X_8, X_9, respectively) were extracted by principal components analysis with SPSS 18.0 statistical software. Variance analysis of the original test data was presented in Table 2 and the degree of correlation of original variables and the principal component was presented in Table 3.

Table 3: Degree of correlation of original variables and the principal component

	Variable			
Component	PC1	PC2	PC3	PC4
GA	-0.020	0.785	0.546	-0.109
EGC	0.726	0.033	-0.616	-0.190
+C	0.576	0.736	-0.127	-0.263
Caffeine	-0.327	0.738	-0.284	0.501
EGCG	0.902	-0.286	0.143	0.242
EC	0.883	0.400	0.017	0.014
ECG	0.945	-0.113	0.154	0.190
Total catechins	0.980	-0.040	-0.156	0.053
Theanine	0.835	-0.167	0.361	-0.014

Table 2 showed that cumulative of variance reached 94.862% of the top four principle components, by the principle of cumulative contribution rate of more than 85%, PC1, PC2, PC3 and PC4 were extracted as the factors which reflected much information of primitive variables. The relationship between original variables and the principal components was as following:

$$PC1 = -0.020X_1+0.726X_2+0.576X_3-0.327X_4+0.902X_5+0.883X_6+0.945X_7+0.980X_8+0.835X_9$$
$$PC2 = 0.785X_1+0.033X_2+0.736X_3+0.738X_4-0.286X_5+0.400X_6-0.113X_7-0.040X_8-0.167X_9$$
$$PC3 = 0.546X_1-0.616X_2-0.127X_3-0.284X_4+0.143X_5+0.017X_6+0.154X_7-0.156X_8+0.361X_9$$
$$PC4 = -0.109X_1-0.190X_2-0.263X_3+0.501X_4+0.242X_5+0.014X_6+0.190X_7+0.053X_8-0.014X_9$$

Four principle factors (Total Catechins, GA, EGC, Caffeine) corresponding to local maximum weighing coefficients were selected, which were shown in Table 3. The scatter plot of PC1 vs. PC2 was shown in Fig. 2. All the tea is closely clustered in the each region of the PCs space, but the Yunnan Pu'er and Guangxi Liubao are partly mixed together.

Table 4: Result of classification from different regions

			Predicted group membership					
		Region	1	2	3	4	5	Total
Original	Count	1	18	0	0	0	0	18
		2	0	4	0	0	0	4
		3	0	0	6	0	0	6
		4	0	0	0	5	0	5
		5	0	0	0	0	5	5
	%	1	100.0	0	0	0	0	100.0
		2	0	100.0	0	0	0	100.0
		3	0	0	100.0	0	0	100.0
		4	0	0	0	100.0	0	100.0
		5	0	0	0	0	100.0	100.0
Cross-validated	Count	1	18	0	0	0	0	18
		2	0	4	0	0	0	4
		3	1	0	5	0	0	6
		4	0	0	0	4	1	5
		5	0	0	0	1	4	5
	%	1	100.0	0	0	0	0	100.0
		2	0	100.0	0	0	0	100.0
		3	16.7	0	83.3	0	0	100.0
		4	0	0	0	80.0	20.0	100.0
		5	0	0	0	20.0	80.0	100.0

1: Yunnan Pu'er tea; 2: Hunan Fuzhuan tea; 3-Guangxi Liubao tea; 4: Hubei Qingzhuan tea; 5: Sichuan Kangzhuan tea

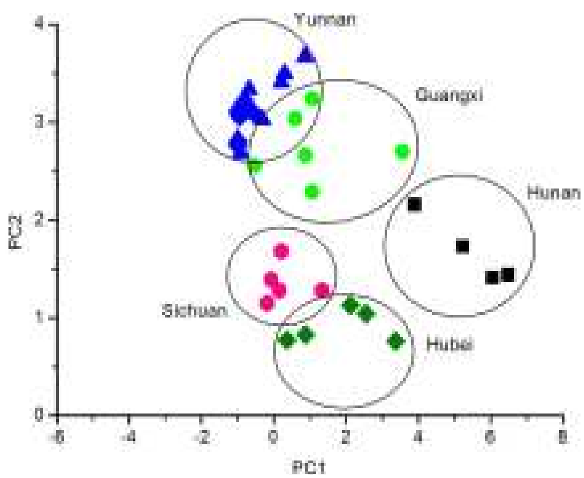

Fig. 2: Results of scattered points for the first and second principal component of dark teas

Results of discriminate analysis: Four principal components had been extracted by principal components analysis and the discriminate functions from the extracted chemical factors were established as the following equations:

Yunnan = 24.577 *Total catechins+78.093* Caffeine-50.554 *EGC-2.021 *GA-143.132
Hunan = 58.963 *Total catechins+63.896 *Caffeine-88.161 *EGC-26.121 *GA-135.104
Guangxi = 22.424 *Total catechins+71.360* Caffeine-34.355 *EGC-9.321 *GA-123.236
Hubei = 17.311 *Totalcatechins+28.298* Caffeine-23.861 *EGC-7.098*GA-23.287
Sichuan = 13.237 *Total catechins+35.205* Caffeine-24.862 *EGC-0.683 *GA-30.827

The values of above functions were calculated by the content of total catechins, caffeine, EGC and GA, respectively. Region corresponding to the maximun value was the unknown tea sample belonged to Wentong and Jie (2007).

Table 4 showed the classification result of five regions. It was found that the discriminate rate was 100%, while the cross-validated rate was 92.1%, which demonstrated that Bayes discriminate analysis had high and robust classification performance.

CONLUSION

HPLC was utilized to measure main chemical composition simultaneously in samples from five dark tea origins. The tea sample contents were used as chemical descriptors to identify the origins of dark tea. Bayes analysis was applied as a pattern recognition tool to develop identification models. The overall results sufficiently demonstrated that it was feasible to realize identification of dark tea Origins according to chemical composition combined with Bayes Classification Pattern Recognition. We conclude that Bayes pattern recognition can be applied to quantitatively identify the dark tea origins based on their chemical component.

ACKNOWLEDGMENT

This study has been financially supported by the special public welfare research of the general administration of quality supervision, inspection and quarantine of P.R.C (Grant No. 201410225) the Natural Science Foundation in Anhui Province (Grant No. 1408085MC61).

Conflict of interest: We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work.

REFERENCES

Duda, R.O., 2003. Pattern Classification. 2nd Edn., China Machine Press, Beijing.

Fulin, Y., D. Fangming and Z. Lingyan, 2006. Development of black tea microbiology. J. Microbiol., 26(1): 81-84.

He, W., X. Hu, L. Zhao, X. Liao, Y. Zhang, M. Zhang and J. Wu, 2009. Evaluation of Chinese tea by the electronic tongue: Correlation with sensory properties and classification according to geographical origin and grade level. Food Res. Int., 42(10): 1462-1467.

ISO 10727:2002. Tea and Instant Tea in Solid Form -- Determination of Caffeine Content -- Method using High-Performance Liquid Chromatography. Retrieved from: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=31494.

ISO 14502-2:2005. Determination of Substances Characteristic of Green and Black Tea -- Part 2: Content of Catechins in Green Tea -- Method using High-Performance Liquid Chromatography. Retrieved from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=31357.

ISO/WD19563. Determination of Theanine in Tea Using High-performance Liquid Chromatography.

JingMing, N., Z. ZhengZhu, G. XunGang, W. XiaoChun and S. WenTong, 2010. Fingerprint identification method of Pu’er raw tea based on high performance liquid chromatography profiles. T. Chinese Soc. Agric. Eng., 26(3): 243-248.

Kodama, S., Y. Ito, H. Nagase, T. Yamashita, T. Kemmei, A. Yamamoto and K. Hayakawa, 2007. Usefulness of catechins and caffeine profiles to determine growing areas of green tea leaves of a single variety, Yabukita, in Japan. J. Health Sci., 53(4): 491-495.

Shushen, L. and Y. Zhongsheng, 1999. Basic Chemometrics. Science Press, Beijing.

Wan, X.C., 2003. Tea Biochemistry. 3rd Edn., China Agriculture Press, Beijing, China, pp: 254-275.

Wei, Y., 2011. The application of the chemical factors of tea quality to the discrimination between oolong teas and white teas. Anhui Agriculture University, Heifei.

Wei-ping, S., L. Hong and Z. Wu-jiu, 2007. Main components analysis method & its application in the evaluation of beer flavor. Liquor-Making Sci. Technol., 11: 107-110.

Wentong, Z. and Y. Jie, 2007. SPSS Statistical Analysis Basic Tutorial. Higher Education Press.

Wold, S., K. Esbensen and P. Geladi, 1987. Principle component analysis. Chemometr. Intell. Lab., 2: 37-52.

Xin-He, Y., W. Li-Li, H. Jian-An *et al*., 2012. Fourier transform infrared spectroscopic analysis of five samples of dark tea. Food Sci., 33(14): 203-205.

Yan, H. and Q. Gui-nian, 2014. Study on the FTIR fingerprinting of dark teas from different regions in China. J. Nucl. Agr. Sci., 28(4): 684-691.

Zengsheng, W., S. Zhaopeng, L. Zhonghua, H. Jian'an, W. Qiongying and Z. Shangtong, 1991. Discussion on the mechanism of quality and flavour formation of dark green tea. J. Tea Sci., 11(1): 1-9.