Application of FT-NIR spectroscopy for simultaneous estimation of taste quality and taste-related compounds content of black tea

Quansheng Chen[1] · Min Chen[1] · Yan Liu[1] · Jizhong Wu[1] · Xinyu Wang[1] ·
Qin Ouyang[1] · Xiaohong Chen[2]

Abstract Fourier transform near-infrared spectroscopy (FT-NIR) coupled to chemometric algorithms such as back propagation (BP)-AdaBoost and synergy interval partial least square (Si-PLS) were deployed for the rapid prediction taste quality and taste-related components in black tea. Eight main taste-related components were determined via chemical analysis and Pearson correlations. The achieved chemical results of the eight taste-related components in black tea infusion were predicted based on 160 tea samples obtained from different countries. Prediction results revealed BP-AdaBoost models gave superior predictions, with all the correlation coefficients of the prediction set (R_p) > 0.76, and the root mean square error values of the prediction set (RMSEP) < 1.7% compared with Si-PLS models ($0.71 \leq Rp \leq 0.94$, $0.08\% \leq RMSEP \leq 1.73\%$). This implies that FT-NIR combined to BP-AdaBoostis capable of being deployed for the rapid evaluation of black tea taste quality and taste-related components content simultaneously.

Keywords Taste quality · Taste-related components · NIR spectroscopy · Multivariate calibration · Black tea

✉ Quansheng Chen
q.s.chen@hotmail.com

1 School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, People's Republic of China

2 Zhenjiang Jinshan Cuiya Tea Industry Co. Ltd, Zhenjiang 212021, People's Republic of China

Introduction

Tea is one of the most pleasant and popular beverages in the world, because its infusion does not only give specific taste, but also provides many essential nutrient compounds for human health promotion (Petta and Marchesini 2017). Generally, tea quality is determined by the shape of the leaf, the colour of the liquid, the aroma and taste. Among all the factors, taste quality plays a decisive role in consumers' preferences. Traditionally, taste quality of black tea is evaluated by the tea experts through sensory analysis (Zhu et al. 2017). Although experts can describe various taste attributes and evaluation scores could be influenced by fatigue. In addition, it is difficult to establish standard terms that can describe tea taste among tea researchers from different country. If taste quality could be related to a specific compounds or group of compounds, then such indices may become standard for describing tea quality among different laboratories or cultures. However, there is currently no research available on the simultaneous estimation of taste quality and taste-related compounds content.

The processing steps of black tea involve harvesting, withering, leaf crushing, fermentation and drying and sieving (Vargas and Vecchietti 2016). However, the unique taste of black tea forms only during fermentation. In this process, catechins are chemically transformed to a number of oxidation products, such as total polyphenols (TPs), free amino acids, theanine, caffeine, catechins and theaflavins (TFs), which contribute largely to the taste of black tea (Kraujalytė et al. 2016; Li et al. 2013). The content of these biochemical compounds are normally detected by chemical analyses which are time-consuming, tedious, expensive and involves cumbersome operations.

Recently, FT-NIR analytical technique had proven to be a powerful analytical tool because of its low-cost and nondestructiveness. It can provide much information about chemical composition through analysis of molecular bonds (e.g. C–H, O–H, N–H and S–H). Owing to the great progress in the data modeling, multivariate statistical techniques can be used to analyse relationships between complex FT-NIR spectroscopy data matrices and sensory score matrices or physicochemical indices. Therefore, FT-NIR spectroscopy can be a useful tool for rapid estimation of the food quality.

This study therefore seeks to attempt the identification of compounds related to taste via the correlation of chemical compositions and infusion taste data; and establishment of models to estimate the taste quality and taste-related compounds content of black tea using FT-NIR.

Materials and methods

Materials

A total of 160 samples of black tea leaves were collected from Burundi, China, Indonesia, India, Kenya, Malawi, Mozambique, Myanmar, Rwanda, Sri Lanka, Tanzania, Uganda and Vietnam; and these were stored in a refrigerator at 4 °C. Distilled water was used throughout the experiment. Standard amino acids, caffeine, individual theaflavins [theaflavin (TF1), theaflavin-3-gallate (TF2A), theaflavin-3'-gallate (TF2B), theaflavin-3,3'-digallate (TF3)] and individual catechins [catechin (C), epigallocatechin gallate (EGCG), epicatechin gallate (ECG), epicatechin (EC), epigallocatechin (EGC)] used in the high performance liquid chromatography (HPLC) were purchased from Sigma Aldrich (Shanghai, China) Trading Company Limited.

Determination of biochemical components

Water extract was determined by convention standard method (ISO 9768-1994) whereas the concentrations GA, caffeine, C, EC, EGC, EGC, EGCG, ECG and total catechin were determined by standard method (ISO 14502-2). Each sample was quantified by spectrophotometer at 570 nm wavelength using theanine as standard reference after extraction, filtration (de-compressional filtration), and then cooled to room temperature (Wang et al. 2010). Theanine was determined by HPLC described by Ying et al. (2005). Polyphenols content were determined according to the standard method (ISO 14502-1); whereas theaflavins, which have been reported to contribute to the astringency of black tea, were determined according to method reported in (Owuor and Obanda 1995).

Taste quality evaluation

Samples were subjected to a taste evaluation according to the National Standard of China (GB/T 23776-2009). The evaluation was conducted in the sensory laboratory of the Department of Tea Science at Anhui Agriculture University of China at room temperature (about 25 °C). Evaluation result of taste quality was scored based on a 10-point scale by 9 professional evaluators with evaluations done three times on different days.

FT-NIR spectral collecting and preprocessing

The raw NIR spectra were acquired by Bruker MPA Fourier transform NIR spectrometer (Bruker Optik GmbH, Ettlingen, Germany) based on a diffuse reflectance mode. 160 samples were ground, sieved with a 500 μm mesh, transferred into quartz cuvette and their spectra (32 scans per sample) acquired within l range of 12,500–4000 cm^{-1}. Background and noise in raw spectra were corrected via standard normal variate (SNV) transformation.

Data analysis

Pearson correlation coefficient was used to determine the degree of correlation between each chemical components and taste score in SPSS 16.0 (Windows 7) after which the synergy interval partial least squares (Si-PLS) and BP-AdaBoost were used to build linear and nonlinear models respectively for the prediction of the taste quality parameters stated in this work. Si-PLS and BP-Adaboost modeling procedures can be found in Online Resource 1.

The performances of these two models were evaluated based on the magnitudes of their correlation coefficients of calibration (R_c) and prediction (R_P) sets; and root mean square error in the calibration (RMSEC) and prediction (RMSEP) sets. Higher R_c and R_p; and smaller RMSEC and RMSEP are indicative of a robust model. Moreover, the differences between these performance indicators must be small, with $R_c > R_p$ and RMSEC < RMSEP. All data processing and analyses were executed using Matlab Version 7.10.0 (Math Works, Natick, USA) in Microsoft Windows 7.

Results and discussion

Correlation analysis between the taste score and chemical composition

Pearson correlation coefficients were used to determine the relationship between the taste score and chemical composition. *P*-values $< \propto = 1\%$ were considered statistically

significant. Based on calculation, a strong correlations was found between taste score and (water extracts, r = 0.738; caffeine, r = 0.691; free amino acids, r = 0.613; tea polyphenols (TPs), r = 0.556; TF2B, r = 0.535; TFs, r = 0.461; TF2A, r = 0.453 and total catechins, r = 0.429). On the other hand, there were weaker correlations between taste score (gallic acid, r = 0.125; TF3, r = 0.114; (−)-epicatechin (EC), r = 0.247; (−)-epigallocatechin-3-gallate (EGCG), r = 0.291 and TF1, r = 0.308). These correlation results suggest the water extracts, caffeine, free amino acids, TPs, catechins, TFs, TF2B and TF2A present in the black tea were the main contributors to the measured taste quality. These concur with previous reports on the taste-related compounds in tea (Ahmed et al. 2010; Yu et al. 2014).

TPs contribute mainly to bitter and astringent taste, whereas the free amino acids account for the umami taste (Kaneko et al. 2006). During the fermentation period, the combination of free amino and ortho-quinones accounts for the aroma change from bitter to sweet. Similarly, theanine, as the most abundant amino acid, was degraded and thus contributed less to the taste score (Sharma et al. 2015). Caffeine is the major contributor to bitterness in tea. TFs are characteristic metabolites in black tea which are formed from the dimerization of catechins due to enzymatic oxidations during tea fermentation process. Aqueous solutions of TFs isolated from black tea have been reported to exhibit astringent taste close to the sensation given by black tea liquor (Millin et al. 1969). Among the major TFs, TF2A and TF2B showed very low contents in the tea leaves initially, but increased rapidly during the fermentation process, in which enzymatic oxidation of catechins occurred (Tan et al. 2016). Thus, based on these relations, the contents prediction of all the afore-stated taste-related components are expected to provide the chemometrics foundation for taste sensing.

Pretreatment of FT-NIR spectra

FT-NIR spectra of 160 black tea samples were collected in the full-spectrum range of 12,500–4000 cm^{-1} with the raw and SNV preprocessed spectra shown respectively in Online Resource 2 (Fig. S1a, b). As seen from Fig. S1b, all the samples exhibited a similar trend in absorbance, but their NIR spectra were influenced by orientation, surface characteristics, and most importantly their biochemical constituents (Sunoj et al. 2016). Peaks at 8200–8600, 6700–7000, 5600–6000, 5000–5300, 4400–4800 and 4200–4400 cm^{-1} occurred due to the vibrations of 2nd overtones of C–H groups; 1st overtones of C–H groups; C–H first overtone and S–H first overtone; 2nd overtones of C=O groups; and combinations of amides and C–H stretching (CH$_2$ groups) in the samples respectively as

reported by Cai et al. (2015). These spectra provide chemical structure information about the samples and suggest NIR could be used for rapid sensing of black tea taste quality.

Quantitative analysis of Si-PLS models and BP-AdaBoost models

Calibration of models

Online Resource 3 (Table S1) shows the reference measurement results by chemical analysis and taste evaluation scores both in calibration and prediction sets. It can be seen that each range of reference results in prediction are contained in the calibration set (prediction set < calibration set).

Si-PLS models

Based on optimal Si-PLS model, the associated chemical information corresponding to the intervals selected for the contents of polyphenol, water extracts, free amino acids, caffeine, total catechin, TF2A, TF2B and theaflavins are summarized in Table 1.

To test the performance of each Si-PLS model, external validation was applied to verify its applicability and accuracy using the prediction set of the samples. The best results of the optimal models for the determination of taste scores based on the caffeine, total polyphenol, water extracts, free amino acid contents, caffeine, total catechins, TF2A, TF2B and theaflavins contents of tea are seen in Table 2. It can be seen that all the Si-PLS models had R$_P$ in the range of 0.71–0.95 and low RMSEP between 0.085 and 2.35%, when PCs were in the range of 5–10. These results imply each Si-PLS model achieved satisfactory performance.

BP-AdaBoost models

BP-AdaBoost model was attempted for the prediction of the taste scores and taste-related components of black tea. PCs extracted by principal component analysis (PCA) from spectra were used as the inputs of model. PCs numbers and Φ were optimized by cross-validation and the optimum ones selected for the establishment of the models. The best BP-AdaBoost model for taste score prediction was achieved with 5 PCs and Φ = 0.014, yielded (R$_c$ = 0.9278, R$_p$ = 0.8817, RMSEC = 0.4620 and RMSEP = 0.5183). However, the best BP-AdaBoost models achieved for the prediction of tea polyphenol, water extracts, free amino acids, caffeine, total catechins, TF2A, TF2B and theaflavin contents yielded (0.8483 ≤ R$_c$ ≤ 0.9709, 0.7755 ≤ R$_p$ ≤ 0.9656, 0.0664% ≤ RMSEC ≤ 1.2326% and 0.0764% ≤ RMSEP

Table 1 Selected spectra intervals by Si-PLS algorithm and their corresponding bonds

Component	Spectra interval/cm^{-1}	Corresponding bonds
Water extracts	11,359.138–10,796.002	O–H second overtone and C–H third overtone
	6823.197–6260.061	C–H first overtone
	4559.083–3999.805	C–H and C–H combinations and C–H and C–C combinations
Tea polyphenol	10,364.008–9835.586	N–H second overtone
	8234.893–7706.471	C–H first overtone
	4524.370–3999.805	C–H and C–H combinations
Free amino acids	10,225.152–9662.017	N–H second overtone
	8524.174–7961.039	–CH=CH second overtone
	7957.182–7394.046	First overtone of C–H combinations
	4559.08–3999.805	C–H and C–H combinations and C–H and C–C combinations
Caffeine	7390.189–6827.054	N–H first overtone
	6256.204–5693.068	C–H first overtone and S–H first overtone
	4559.083–3999.805	C–H + C–H combinations and C–H and C–C combinations
Total catechins	9091.167–8528.032	C–H second overtone
	6256.204–5693.068	C–H first overtone and S–H first overtone
	4559.083–3999.805	C–H and C–H combinations and C–H and C–C combinations
TF2A	10,055.440–9449.877	N–H second overtone
	5812.638–5210.932	C=O stretch second overtone
	4601.511–3999.805	C–H and C–H combinations, C–H and C–C combinations, N–H and O–H combinations
TF2B	9658.160–9095.024	Water absorption
	8524.174–7961.039	–CH=CH second overtone
	5689.211–5126.076	O–H combination and C–H first overtone
	4559.083–3999.805	C–H and C–H combinations and C–H and C–C combinations
Theaflavins	11,359.138–10,796.002	O–H second overtone and C–H third overtone
	10,792.145–10,460.435	Water absorption
	4559.083–3999.805	C–H and C–C combinations

TF2A Theaflavin-3-gallate, *TF2B* theaflavin-3′-gallate (Cai et al. 2015)

Table 2 The best results from the optimal models of taste score and each taste-related component both in the calibration set and the prediction set based on using Si-PLS algorithm

Quality parameters	n	Selected intervals	PCs	Calibration set		Prediction set	
				R$_c$	RMSEC (%)	R$_p$	RMSEP (%)
Taste score	15	[1 14 15]	5	0.8517	0.568	0.8462	0.638
Water extracts	15	[3 11 15]	9	0.9376	1.419	0.9089	1.73
Total polyphenols	16	[5 9 16]	6	0.734	2.29	0.7126	2.35
Free amino acids	15	[5 8 9 15]	9	0.9373	0.246	0.8866	0.357
Caffeine	15	[10 12 15]	9	0.9517	0.19	0.9428	0.205
Total catechins	15	[7 12 15]	10	0.7794	0.821	0.7199	0.947
TF2A	14	[5 12 14]	10	0.728	0.149	0.7194	0.178
TF2B	15	[6 8 13 15]	10	0.8324	0.070	0.8074	0.0805
Theaflavins	15	[3 4 15]	8	0.7987	0.296	0.7726	0.34

n Number of intervals, *PCs* principal components number, *R$_c$* correlation coefficient of calibration, *RMSEC* root mean square error in the calibration set, *R$_p$* correlation coefficient of prediction, *RMSEP* root mean square error values of the prediction set, *TF2A* theaflavin-3-gallate, *TF2B* theaflavin-3′-gallate

≤ 1.6965%). All the models have optimized PCs in range of 5–9, R$_P$ greater than 0.76, and RMSEP lower than 1.7%. The individual model results and the various PCs with which they were achieved are summarized in Online Resource 4 (Table S2).
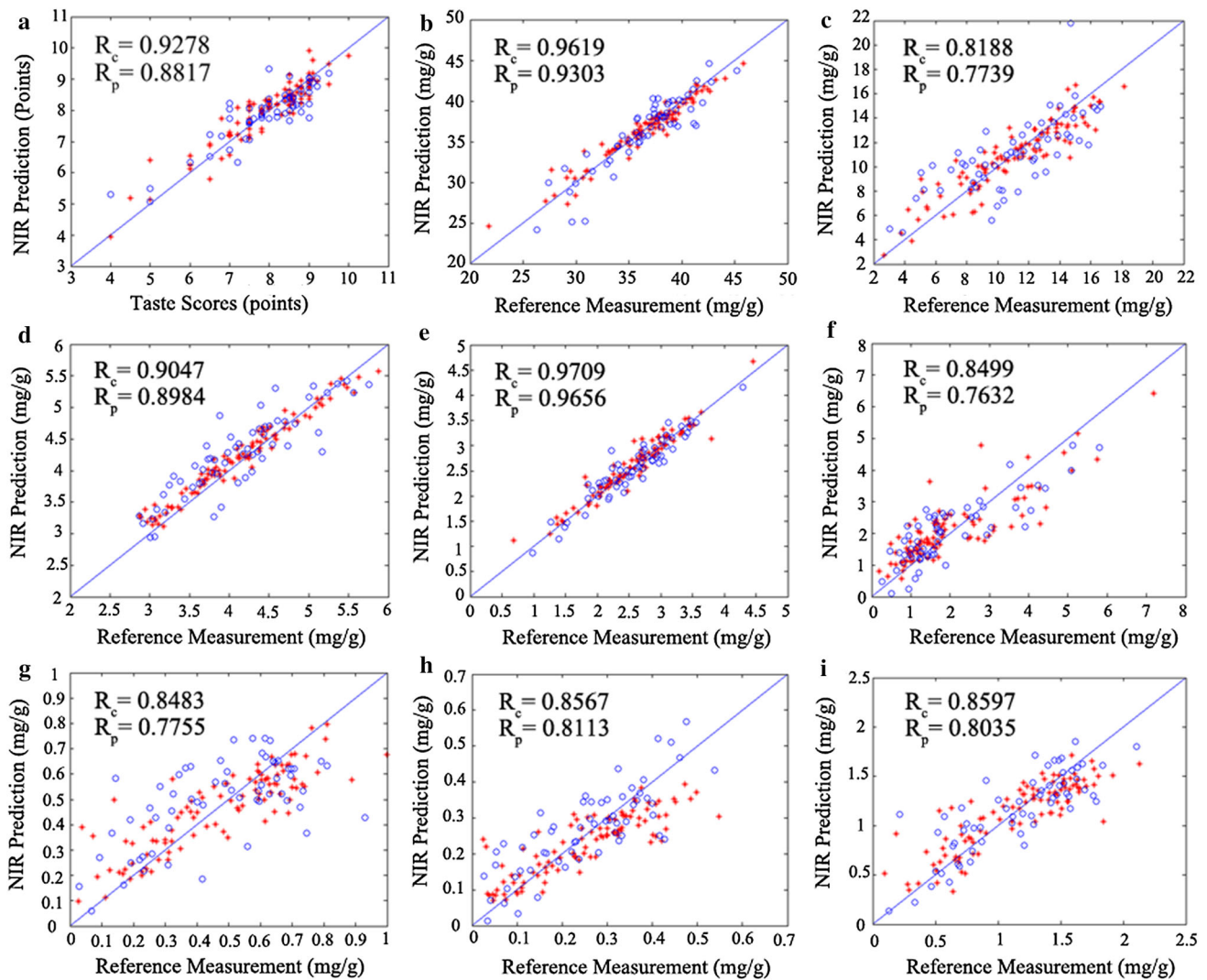
Fig. 1 Scatters plots of each BP-AdaBoost model showing the correlation between reference measurement and NIR prediction in calibration set (*) and the prediction set (o): **a** taste score model, **b** water extracts model, **c** tea polyphenol model, and **d** free amino acids model, **e** caffeine model, **f** total catechins model, **g** TF2A model, **h** TF2B model, **i** theaflavins model

Results of Si-PLS and BP-AdaBoost models proved both can be used to evaluate the taste scores and taste-related components content of black tea.

For Si-PLS, the full spectrum was split into varying number of intervals and then specific intervals were combined as the effective variables for modeling. However, only the spectra data collected within 3–4 intervals (as seen in Table 2) were used to predict related contents. Black tea fermentation is a sophisticated process, in which numerous reactions take place simultaneously. It involves series of complicated enzymatic oxidation of tea polyphenols; thus, the correlation between the NIR spectra and key quality parameters in black tea infusion could incline to nonlinearity, in which case linear regression tools may prove incapable of providing a complete solution for the prediction of complex nonlinear key quality relationships in this

work. In contrast, BP-AdaBoost shows superiority, since it has self-learning and self-adjustment ability. BP-AdaBoost was therefore applied to evaluate the taste scores and taste-related components content of black tea using FT-NIR spectroscopy. The scatter plots between reference values and NIR predicted values both in the calibration and prediction sets by BP-AdaBoost model are shown in Fig. 1. It is evident that each model has a high correlation coefficient within both the calibration and prediction sets.

Conclusion

TPs, water extracts, free amino acids, water extracts, caffeine, TF2B, TF2A, TF, catechins were determined to be the main contributors to the taste quality of black tea

through Pearson correlation analysis. Combined with BP-AdaBoost algorithm, FT-NIR spectroscopy system showed good performance in the rapid prediction of the taste scores and taste-related components content of black tea. Based on the results of this study, FT-NIR spectroscopy proved a potential technique for the rapid sensing of taste quality and taste-related components contents simultaneously. This outcome provides an effective tool for effective quality control, process monitoring and produce pricing in the tea industry.

References

Ahmed S, Unachukwu U, Stepp JR, Peters CM, Long C, Kennelly E (2010) Pu-erh tea tasting in Yunnan, China: correlation of drinkers' perceptions to phytochemistry. J Ethnopharmacol 132:176–185. https://doi.org/10.1016/j.jep.2010.08.016

Cai JX, Wang YF, Xi XG, Li H, Wei XL (2015) Using FTIR spectra and pattern recognition for discrimination of tea varieties. Int J Biol Macromol 78:439–446. https://doi.org/10.1016/j.ijbiomac.2015.03.025

Kaneko S, Kumazawa K, Masuda H, Henze A, Hofmann T (2006) Molecular and sensory studies on the umami taste of Japanese green tea. J Agric Food Chem 54:2688

Kraujalytė V, Pelvan E, Alasalvar C (2016) Volatile compounds and sensory characteristics of various instant teas produced from black tea. Food Chem 194:864–872. https://doi.org/10.1016/j.foodchem.2015.08.051

Li S, Lo CY, Pan MH, Lai CS, Ho CT (2013) Black tea: chemical analysis and stability. Food Funct 4:10–18. https://doi.org/10.1039/C2FO30093A

Millin DJ, Crispin DJ, Swaine D (1969) Nonvolatile components of black tea and their contribution to the character of the beverage. J Agric Food Chem 17:717–722. https://doi.org/10.1021/jf60164a026

Owuor PO, Obanda M (1995) Clonal variation in the individual theaflavin levels and their impact on astringency and sensory evaluations. Food Chem 54:273–277. https://doi.org/10.1016/0308-8146(95)00046-L

Petta S, Marchesini G (2017) Corrigendum to "Coffee and tea breaks for liver health" [J Hepatol 67 (2017) 221–223]. J Hepatol 67:1122. https://doi.org/10.1016/j.jhep.2017.07.016

Sharma P et al (2015) Monitoring the fermentation process of black tea using QCM sensor based electronic nose. Sens Actuators B Chem 219:146–157. https://doi.org/10.1016/j.snb.2015.05.013

Sunoj S, Igathinathane C, Visvanathan R (2016) Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. Comput Electron Agric 124:234–242. https://doi.org/10.1016/j.compag.2016.04.012

Tan J et al (2016) Study of the dynamic changes in the non-volatile chemical constituents of black tea during fermentation processing by a non-targeted metabolomics approach. Food Res Int 79:106–113. https://doi.org/10.1016/j.foodres.2015.11.018

Vargas R, Vecchietti A (2016) Influence of raw material moisture on the synthesis of black tea production process. J Food Eng 173:76–84. https://doi.org/10.1016/j.jfoodeng.2015.10.043

Wang K et al (2010) Analysis of chemical components in oolong tea in relation to perceived quality. Int J Food Sci Technol 45:913–920. https://doi.org/10.1111/j.1365-2621.2010.02224.x

Ying Y, Ho JW, Chen ZY, Wang J (2005) Analysis of theanine in tea leaves by HPLC with fluorescence detection. J Liq Chromatogr Relat Technol 28:727–737. https://doi.org/10.1081/JLC-200048894

Yu P, Yeo ASL, Low MY, Zhou W (2014) Identifying key non-volatile compounds in ready-to-drink green tea and their impact on taste profile. Food Chem 155:9–16. https://doi.org/10.1016/j.foodchem.2014.01.046

Zhu H, Ye Y, He H, Dong C (2017) Evaluation of green tea sensory quality via process characteristics and image information. Food Bioprod Process 102:116–122. https://doi.org/10.1016/j.fbp.2016.12.004