JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

High-Throughput Technique for Comprehensive Analysis of Japanese Green Tea Quality Assessment Using Ultra-performance Liquid Chromatography with Time-of-Flight Mass Spectrometry (UPLC/TOF MS)

Wipawee Pongsuwan,[†] Takeshi Bamba,[†] Kazuo Harada,[†]
Tsutomu Yonetani,[‡] Akio Kobayashi,[†] and Eiichiro Fukusaki*,[†]

Department of Biotechnology, Graduate School of Engineering, Osaka University,
2-1, Yamadaoka, Suita, Osaka 565-0871, Japan, and Tea Branch, Nara Prefecture Agricultural
Experiment Station, Otsu 470-1, Yadawara, Nara 630-2166, Japan

Applications of metabolomics techniques along with chemometrics provide an understanding in the relationship between metabolome of green tea and its quality. A coupled of ultra-performance liquid chromatography with time-of-flight mass spectrometry (UPLC/TOF MS) allowed a high-throughput and comprehensive analysis with minimal sample preparation. Using this technique, a wide range of metabolites were investigated. Data analysis was rapid, considering that the fingerprinting technique was performed. A set of green tea samples from 2006 tea contest of the Kansai area was analyzed to prove usefulness of the developed technique. Green tea with different qualities were discriminated through principal component analysis (PCA). Consequently, projection to latent structure by means of partial least-squares (PLS) was performed to create a constructive quality-predictive model by means of metabolic fingerprinting. Beside epigallocatechin, other predominant catechins, including epigallocatechin gallate and epicatechin gallate, detected in green tea were found to be significant biomarkers to the high quality of Japanese green tea (Sencha).

KEYWORDS: UPLC/TOF MS; metabolic fingerprinting; quality of green tea; prediction model

INTRODUCTION

Sensory evaluation of green tea has traditionally been assessed by highly trained specialists, who evaluate product quality based on leaves appearance, aroma, and taste of brew. The sweet, brothy taste of green tea originates from amino acids, especially theanine, which accounts for about two-thirds of the total amino acids content in tea leaves (*1*). In addition, caffeine and catechins have generally been considered as components responsible for the characteristic astringent and bitter taste of tea brews (*2*). Recently, many analytical methods and instrumental measurements have been applied to explore the correlation between the quality of green tea and its chemical constituents. Each analytical technique was developed for the characterization on the group of interested compounds with regard to quality (*3−5*) or antioxidant activities (*6−8*). However, correlation and balance of compounds are also important to assess quality of green tea.

Therefore, techniques in metabolomics were applied to acquire the relationship between green tea components and their quality. The key of metabolomics is comprehensive chemical analysis of metabolites and the computation of huge data sets. Metabolomics is principally required to determine all metabolites in plant extracts; however, no single technology is available. This is because the plant is a rich source of diverse functional biochemistry with wide concentration ranges of compounds. Many analytical methods were developed and introduced to metabolomics analysis, such as mass spectrometry (MS) (*9, 10*), nuclear magnetic resonance (NMR) (*11, 12*), and infrared (IR) and Raman spectroscopy (*13, 14*). Because of its high sensitivity, which is perhaps the most important requirement for metabolomics, MS has established itself as a method of choice in plant research (*15*). The resolution and selectivity of this technology can be enhanced or modified by coupling with gas chromatography (GC) or liquid chromatography (LC). Normally, the instruments selection depends upon the group and type of compounds to be analyzed. GC/MS allows for the identification, quantification, and also structural elucidation of metabolites. However, there are some limitations to consider. This approach can only be used with volatile compounds or compounds that can be volatile by derivatization. In contrast, LC/MS can be

* To whom correspondence should be addressed: Department of Biotechnology, Graduate School of Engineering, Faculty of Engineering, Osaka University, 2-1, Yamadaoka, Suita, Osaka 565-0871, Japan. Telephone/Fax: +81-6-6879-7424. E-mail: fukusaki@bio.eng.osaka-u.ac.jp.
† Osaka University.
‡ Nara Prefecture Agricultural Experiment Station.

adapted to a wider range of molecules, including secondary metabolites, such as alkaloids, flavonoids, glucosinolates, isoprenes, polypropanoids, saponins, and carotenoids (*16*). In addition, LC/MS operates at lower analysis temperature compared to GC/MS, which enables the analysis of heat-labile metabolites that are commonly degraded during GC analysis, and it also does not require sample derivatization, which simplified the sample preparation steps (*17*).

Metabolic fingerprinting was the selected approach for our study. It is a technique that involves the classification of samples based on their biological background rather than characterizing a limited number of individual compounds. With the fingerprinting technique, the assigned problem of each metabolite could be ignored. Fingerprinting requires an analytical technique that can handle a large number of samples with minimal sample preparation but is still capable of providing relevant chemical fingerprinting (*18*). By improving the LC technology, the newest technique of ultra-high-performance liquid chromatography (UPLC) has been developed. UPLC using 1.7 μm particles and a properly holistically designed system provide significantly more resolution while reducing run times and improve sensitivity for the analyses of many compound types. UPLC also presents the possibility to extend and expand the utility of chromatography compared to the conventional technique, such as high-performance LC (HPLC). An ultra-high pressure system allows for the usage of small-particle-packed columns with small diameter, which gives a positive effect on both system efficiency and analysis time. A flow was increased to 3 times because of smaller particles, and shortening of the column by one-third made separation complete in $^1/_9$ time while maintaining resolution. UPLC provides significant advantages concerning selectivity, sensitivity, and speed. The technique is suitable for fingerprinting analysis, considering its speed, robustness, and high sample throughput. It also allows for detection of wide-range metabolites, including both hydrophilic and hydrophobic metabolites within a single chromatographic run (*9*).

In this study, the method for Japanese green tea quality evaluation was developed. From a coupled of UPLC with electrospray ionization mass spectrometry (ESI−MS), a fingerprinting of multiple components of tea was established. Chemometrics analysis was performed to compare signals derived from metabolites present in various grade green tea samples. Consequently, peak identification of selected variables was achieved to reveal significant metabolites for green tea quality determination.

MATERIALS AND METHODS

Materials. The dried leaves after processing of 56 ranked first crop tea samples (spring-harvested, called "Ichi-ban-cha" in Japanese) from the 2006 contest were analyzed. These tea samples enrolled in a commercial tea contest among Kansai area teas were obtained from the Tea Branch of the Nara Prefecture Agricultural Experiment Station. The ranking was determined by the total scores of the sensory tests, which included leaf appearance, smell and color of the brew, and its taste, judged by professional tea tasters.

Sample Preparation for UPLC/MS Analysis. Dried tea leaves (10 mg) in a 2 mL Eppendorf tube were freeze-dried and ground with a Retsch ball mill (20 Hz., 1 min). Samples were extracted by 1 mL of solvent mixture of MeOH, H_2O, and $CHCl_3$ in a ratio of 2.5:1:1 (v/v/v), respectively. The mixture was incubated at 37 °C for 30 min and then centrifuged at 16000*g* and 4 °C for 10 min. Subsequently, 200 μL of the supernatant was transferred to a 1.5 mL Eppendorf tube capped with a pierced cap. The extract was dried in a vacuum centrifuge dryer and freeze-dryer until dry. Afterward, 200 μL of methanol and water containing 0.1% formic acid in a ratio of 4:1 was added. The mixture was filtered through 0.2 μm PTFE filter. The injection volume was 4 μL.
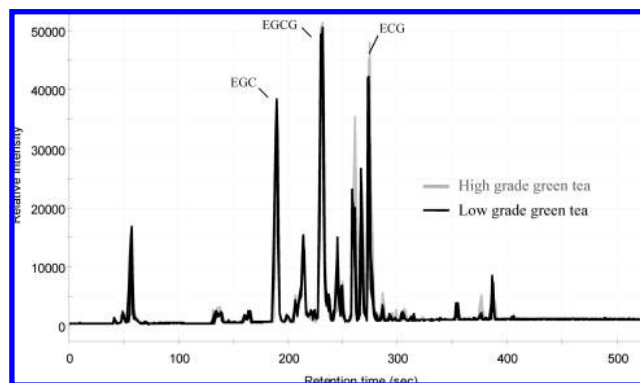

Figure 1. Total ion chromatograms of Japanese green tea dried leaves (high- and low-grade green teas). Key compounds for quality determination of green tea: (−)-epigallocatechin (EGC), (−)-epigallocatechin gallate (EGCG), and (−)-epicatechin gallate (ECG).

UPLC/MS Analysis. The Ultra-performance liquid chromatography (UPLC) was performed on a Waters ACQUITY UPLC system (Waters, Milford, MA), which was equipped with a binary solvent delivery manager, column manager, and sample manager. Detection was performed on a Waters LCT premier XE mass spectrometry (Waters, Milford, MA). The instrument was fitted with an Acquity UPLC BEH C18 column, 1.7 μm, 2.1 × 150 mm (Waters, Milford, MA), operated at 40 °C.

Mobile-Phase Conditions. Linear gradient analysis with mobile-phase A, H_2O (0.1% formic acid), and mobile-phase B, acetonitrile (0.1% formic acid). Mobile-phase B was increased from 0 to 55.6% over 10 min and to 100% at 10.10 min, with a flow rate of 0.3 mL/min.

MS Conditions. The instrument was operated using an ESI source in negative mode. The acquired *m/z* was from 100 to 1000. The ionization parameters were capillary voltage, 2.0 kV; cone voltage, 0.035 kV; the desolvation gas flow, 500 L/h; the desolvation temperature, 350 °C; and source temperature, 100 °C. Analyses were performed using the lock spray to ensure accuracy and reproducibility; leucine−enkephalin was used as the lock mass (*m/z* 554.261).

Data Treatment. Data processing was performed on both Marker-Lynx (software package provided by Waters) and MZmine (http://mzmine.sourceforge.net/). With MarkerLynx, each mass number was analyzed separately in a search peak. The area of the peak would be given an identity of *m/z* and retention time, which were used as fingerprints. For MZmine, raw chromatographic data were converted into NetCDF format,.cdf) before applying to the software. Peak detection finds the peaks corresponding to the compounds. Alignment aims at matching the corresponding peaks across multiple sample runs. Spectral filtering aims at reducing the complexity of data and removing the noise. The role of normalization is to reduce the systematic error by adjusting the intensities within each sample run (*19*).

Multivariate Analysis. Principle component analysis (PCA), SIM-CA-P version 11.0 (Umetrics, Umeå, Sweden), was initially executed to understand the relationship expressed in terms of similarity or dissimilarity among groups of multivariate data. Projections to latent structures by means of partial least-squares (PLS), SIMCA-P version 11.0, was subsequently performed to create a quality-prediction model.

RESULTS AND DISCUSSION

Metabolite Fingerprinting with PCA. The total ion chromatogram of the green tea sample was shown in **Figure 1**. Both high- and low-grade green teas presented the same chromatographic patterns. However, the differences in the intensities of peaks through variety of tea qualities were investigated. In the analyses of complex mixtures, coeluted compounds are often detected. Operative software was applied to extract information out of those complex data. After chromatograms were processed by MarkerLynx (Waters, Milford, MA), it was found that some data were missing at several ranges of retention time. Because
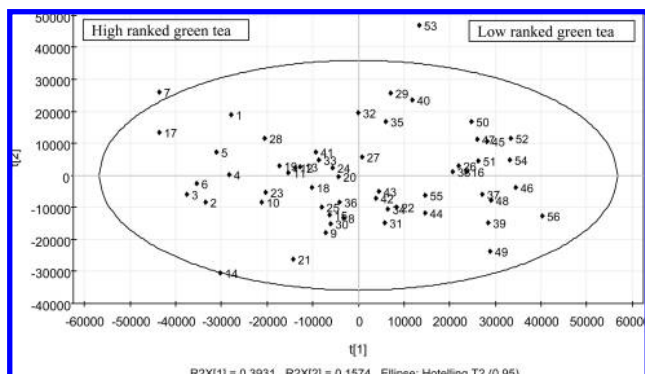
Japanese Green Tea Quality Assessment Using UPLC/TOF MS

J. Agric. Food Chem., Vol. 56, No. 22, 2008   **10707**

Figure 2. PCA shows differences between low- and high-grade green tea samples. High-grade green tea samples were clustered along the negative side of PC1, while low-grade green tea samples were located at the positive side of the PC1 axis.

metabolomics aims to study the entire metabolites and the missing information might have significance to differences in samples, other processing software was applied to complete the information. MZmine is a remarkable choice for processing of LC chromatographic data developed by Katajamaa et al. (*19*). It provides various options for peak detection. Compounds that were first omitted by MarkerLynx were detected with MZmine. Peak detection (*19*) regarding the retention time and *m/z* was performed, followed by filtering to remove the peaks with weakest intensities. In total, 1560 compounds were extracted out of a complex chromatogram.

Ionization was performed on both negative and positive modes. However, chromatographic data from positive mode showed poor stability of ionization. Therefore, multivariate analysis was carried out with data from negative mode only. Unsupervised PCA was applied to evaluate the relationship between green tea samples and different ranking. With eight principal components, explaining 85.0% percent of variation of data, it was found that low- and high-ranked teas were grouped separately along PC axis. High-ranked samples were clustered along the negative side of PC1, while low-ranked samples were gathered more at the positive region (**Figure 2**). This proved that there were differences in metabolite profiling of green tea samples through a variety of product quality. In addition, metabolites that play key roles in differentiating the chromatograms of low- and high-grade green tea were illustrated. Identification of the compounds was performed on the basis of their retention time and accurate mass data, along with confirmation of results from the absorbance spectrum to authentic standards. Dominant compounds were found to be catechins. Especially, (−)-epigallocatechin gallate (EGCG), (−)-epicatechin gallate (ECG), and (−)-epigallocatechin (EGC) were detected as dimer with *m/z* 915.176, 883.181, and 611.142, respectively (**Figure 1**). Apart from EGC, the other compounds were found to be greater in high-grade green tea. However, caffeine, which is a major alkaloid present in tea leaves and also has an effect on product quality, was not detected. This resulted from the selected ionization method (negative mode) because it is not suitable for the detection of compounds belonging to the xanthine group.

Prediction of Quality with PLS. A relationship between the quality of green tea (*y* variables) and its metabolic profiles (*x* variables) was examined. *y* variables were sample-ranking-judged by professional tea tasters, whereas *x* variables were calculated from the peak intensity of each metabolite. A quality-predictive model was constructed by means of PLS regression. Data from all 56 ranked green tea samples were divided into
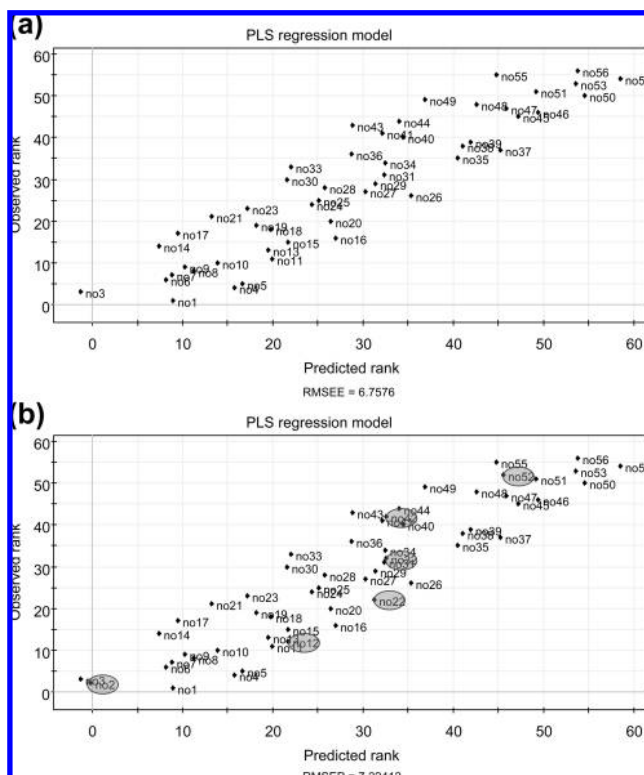

Figure 3. Calibration models with two factors for the prediction of green tea quality (ranking) based on their chromatograms, presented as *y* (observed) versus *y* (predicted), (a) for 50 green tea samples as a training set and (b) for all ranks of both testing (marked by circles) and training sets.

50 training set samples and 6 test set samples by excluding every second sample in every 10th ranked for model validation. The PLS model was built with the remaining data of training set (**Figure 3a**). All variables of X matrix were scaled to center. Two significant components, which was determined by cross-validation were extracted, describing 83.3% of variation in Y (R2Y = 0.833) and predicting 71.4% of variation in Y (Q2 = 0.714). Subsequently, the test set (left out samples) was predicted with this PLS model, resulting in a predictive accuracy for the test samples (RMSEP = 7.22) (**Figure 3b**), compared to the model estimated on the basis of training set samples (RMSEE = 6.75). By demonstrating high values of model goodness (R2Y) and predictivity parameter (Q2), this method is one attractive way for nonspecialist to estimate the quality of green tea and acquire certain quality parameters. In addition, according to the VIP value (the variable influence on projection, *20*), the group of epicatechins (EGCG, ECG, and EGC) along with unidentified compounds with *m/z* 609.088, 687.143, 633.073, and 383.100 were found to be the most influential variables for model construction.

Previously, we have demonstrated two techniques for green tea quality assessment through GC/MS (*20*) and a coupled of pyrolyzer to GC/MS (*21*). The former technique was constructed on the basis of green tea primary metabolites, which was reported to be key compounds for the flavor of the tea. Compound identification could be achieved through the National Institute of Standards and Technology (NIST) library or comparison to authentic standards. However, complicated sample preparation and limitation on the class of compounds was demerit of this approach. The latter technique, PY−GC/MS, expanded the range of detection. All metabolites were investigated in a single analysis. With the functionality of the pyrolyzer, sample preparation, extraction, and

10708 *J. Agric. Food Chem.,* Vol. 56, No. 22, 2008

Pongsuwan et al.

derivatization could be omitted. In this study, the latest developed technique, metabolic fingerprinting through UPLC/TOF MS, provided a rapid analysis with high sample throughput. The technique allows for broad coverage of compounds. From our extraction method, thousand of compounds could be detected within 10 min. This method offered the shortest analysis time with the highest detected number of metabolites among three techniques (*20, 21*). A quality-predictive model showed a high value of R2Y and Q2 with a low prediction error calculated from both test and training sets. Considering that extensive range of compounds was investigated, the acquired information was relatively complete. In addition, the group of epicatechins were identified as significant biomarkers to determine the quality of the product.

LITERATURE CITED

(1) Nakagawa, M. Contribution of green tea constituents to the intensity of taste element of brew. *J. Agric. Food Chem.* **1975**, *22*, 59–64.

(2) Luypaert, J.; Zhang, M. H.; Massart, D. L. Feasibility study for the use of near infrared spectroscopy in the qualitative and quantitative analysis of green tea, *Camellia sinensis* L. *Anal. Chim. Acta* **2003**, *478*, 303–312.

(3) Horie, H.; Kohata, K. Application of capillary electrophoresis to tea quality estimation. *J. Chromatogr., A* **1998**, *802*, 219–223.

(4) Kaneko, S.; Kumazawa, K.; Masuda, H.; Henze, A.; Hofmann, T. Molecular and sensory studies on the umami taste of Japanese green tea. *J. Agric. Food Chem.* **2006**, *54*, 2688–2694.

(5) Yu, H.; Wang, J. Discrimination of LongJing green-tea grade by electronic nose. *Sens. Actuators, B* **2007**, *122*, 134–140.

(6) Tianhong, P.; Joseph, J.; Weidong, L. Potential therapeutic properties of green tea polyphenols in Parkinson's disease. *Drugs Aging* **2003**, *20*, 711–721.

(7) Vagar, M. A.; Hasan, M. Polyphenols from green tea and pomegranate for prevention of prostate cancer. *Free Radical Res.* **2006**, *40*, 1095–1104.

(8) Mustafa, A. V.; Hasan, M. Anti-oxidants from green tea and pomegranate for chemoprevention of prostate cancer. *Mol. Biotechnol.* **2007**, *31*, 52–57.

(9) Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass spectrometry-based metabolomics. *Mass Spectrum. Rev.* **2006**, *26*, 51–78.

(10) Kimberly, R. H.; Amber, H.; Cheolhwan, O.; Xiang, Z.; Jiri, A.; Maria, S. S. Development of GC×GC/TOF−MS metabolomics for use in ecotoxicological studies with invertebrates. *Aquat. Toxicol.* **2008**, *88*, 48–52.

(11) Krishnan, P.; Kruger, N. J.; Ratcliffe, R. G. Metabolite fingerprinting and profiling in plants using NMR. *J. Exp. Bot.* **2004**, *56*, 255–265.

(12) Huifeng, W.; Andrew, D. S.; Adam, H.; Mark, R. V. High-throughput tissue extraction protocol for NMR and MS-based metabolomics. *Anal. Biochem.* **2008**, *37*, 204–212.

(13) Daniel, P. C.; Drew, R. E.; David, J. D.; Timothy, W. C. Raman spectroscopy-based metabolomics for differentiating exposures to triazole fungicides using rat urine. *Anal. Chem.* **2007**, *79*, 7324–7332.

(14) Goodacre, R.; Roberts, L.; Ellis, D. I.; Thorogood, D.; Reader, S. M.; Ougham, H.; King, I. From phenotype to genotype: Whole tissue profiling for plant breeding. *Metabolomics* **2007**, *3*, 489–501.

(15) Roessner, U.; Luedemann, A.; Brust, D.; Fiehn, O.; Linke, T.; Willmitzer, L.; Fernie, A. R. Metabolomic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **2001**, *13*, 11–29.

(16) Fernie, A. R.; Trethewey, R. N.; Krotzky, A. J.; Willmitzer, L. Metabolite profiling: From diagnostics to systems biology. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 1–7.

(17) Bedair, M.; Sumner, L. W. Current and emerging mass-spectrometry technologies for metabolomics. *Trends Anal. Chem.* **2008**, *27* (3), 238–250.

(18) Goodacre, R.; Vaidyanathan, S.; Dunn, W. B.; Harrigan, G. G.; Kell, D. B. Metabolomics by numbers: Acquiring and understanding global metabolite data. *Trends Biotechnol.* **2004**, *22*, 245–252.

(19) Katajamaa, M.; Oresic, M. Processing methods for differential analysis of LC/MS profile data. *BMC Bioinf.* **2005**, *6*, 179.

(20) Pongsuwan, W.; Fukusaki, E.; Bamba, T.; Yonetani, T.; Kobayashi, A. Prediction of Japanese green tea ranking by GC/MS based hydrophilic metabolite fingerprinting. *J. Agric. Food Chem.* **2007**, *55*, 231–236.

(21) Pongsuwan, W.; Bamba, T.; Yonetani, T.; Kobayashi, A.; Fukusaki, E. Quality prediction of Japanese green tea using pyrolyzer coupled GC/MS based metabolic fingerprinting. *J. Agric. Food Chem.* **2008**, *56*, 744–750.