JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Study of Catechin and Xanthine Tea Profiles as Geographical Tracers

PEDRO L. FERNÁNDEZ, FERNANDO PABLOS, MARÍA J. MARTÍN,* AND A. GUSTAVO GONZÁLEZ

Department of Analytical Chemistry, Faculty of Chemistry, University of Seville, 41012 Seville, Spain

The contents of gallic acid, epigallocatechin gallate, epigallocatechin, epicatechin, epicatechin gallate, catechin, caffeine, teophylline, and theobromine were determined in a set of 45 tea samples, including fermented (black and red) and nonfermented (green) teas of different geographical origins (i.e., China, Japan, Kenya, Sri Lanka, and India). A reversed-phase high performance liquid chromatographic method with gradient elution and photometric detection at 275 nm was used to carry out the analysis. Before the HPLC determination, an extraction step was developed using a mixture of acetonitrile and water (60:40, v/v). Pattern recognition techniques involving principal component analysis (PCA) and linear discriminant analysis (LDA) were applied to differentiate the tea samples according to their geographical origins. Catechins, gallic acid, and tea alkaloids are adequate chemical descriptors to distinguish between fermented and nonfermented tea samples cultivated in different geographical areas.

KEYWORDS: Tea; polyphenols; catechins; xanthines; HPLC; chemometrics; pattern recognition

INTRODUCTION

Tea is a beverage consumed in the entire world because of its attractive aroma, taste, and healthy effects. Tea plants (*Camellia sinensis*) are widely cultivated in over 30 countries and are significant in their economies (*1*). Tea-producing areas are located mainly in Southeast Asia, including China, India, Japan, Taiwan, Sri Lanka, and Indonesia, and in central African countries such as Kenya. Nowadays, hundreds of teas are produced. Commercial teas can be generally classified into three major categories: the nonfermented green teas, the partially fermented oolong and paochong teas, and the fully fermented black and pu-erh (red) teas (*2*).

Many factors can constitute important influences on the composition of tea, such as species, season, age of the leaves (plucking position), climate, and horticultural conditions (soil, water, minerals, fertilizers, etc.) (*3*). Catechins, together with phenolic acids such as gallic acid (GA), are a group of polyphenols that constitute up to 30% of the dry weight of the tea leaf and are important factors in the taste of tea. These compounds are the most biologically active group of tea components, especially certain catechins. The major tea catechins are epigallocatechin gallate, epigallocatechin, epicatechin, and catechin (*2*). Some of the biological functions of tea polyphenols are antioxidative activity (*4*), and antihypertensive (*5*) and anticarcinogenic effects (*6*, *7*). Caffeine, theophylline, and theobromine are the main methylxanthines constituting the tea alkaloids, being important factors in the quality of teas. Tea also contains minerals and trace elements such as K, Mn, Cr, Ni, and Zn, which are essential to human health (*8*).

Several studies determining the tea catechins and tea alkaloids, separately (*9*, *10*) and simultaneously (*2*, *11*), have been carried out. Though several authors propose capillary electrophoresis as the technique to be used (*12*−*14*), the analytical method commonly used for the determination of these compounds is high-performance liquid chromatography (HPLC), which currently constitutes the most useful approach for routine analysis. Many works have been reported including HPLC determinations of these tea polyphenols with isocratic (*2*, *15*) and gradient elution (*11*, *16*−*19*).

As it has already been mentioned, climate and agricultural practices, including soil, water, and fertilizers, can be of great influence on the composition of teas. Thus, teas cultivated in different geographical areas will present significant differences in their chemical compositions. Metals have been found to be adequate descriptors to distinguish teas of different geographical origins (*20*, *21*).

In this paper, the contents of five catechins (catechin (C), epicatechin (EC), epigallocatechin (EGC), epicatechin gallate (ECG), and epigallocatechin gallate (EGCG)), one phenolic acid (gallic acid (GA)), and three alkaloids (theobromine (TB), theophylline (TP), and caffeine (CAF)) (**Figure 1**) have been analyzed in teas cultivated in different areas (China, Japan, Kenya, Sri Lanka, and India). These parameters have been considered as chemical descriptors to differentiate teas according to their geographical origins. Pattern recognition techniques,

* To whom correspondence should be addressed (telephone 34-954-557173; fax 34-954-557168; E-mail mmartin@us.es).

10.1021/jf0114435 CCC: $22.00
Published on Web 03/03/2002

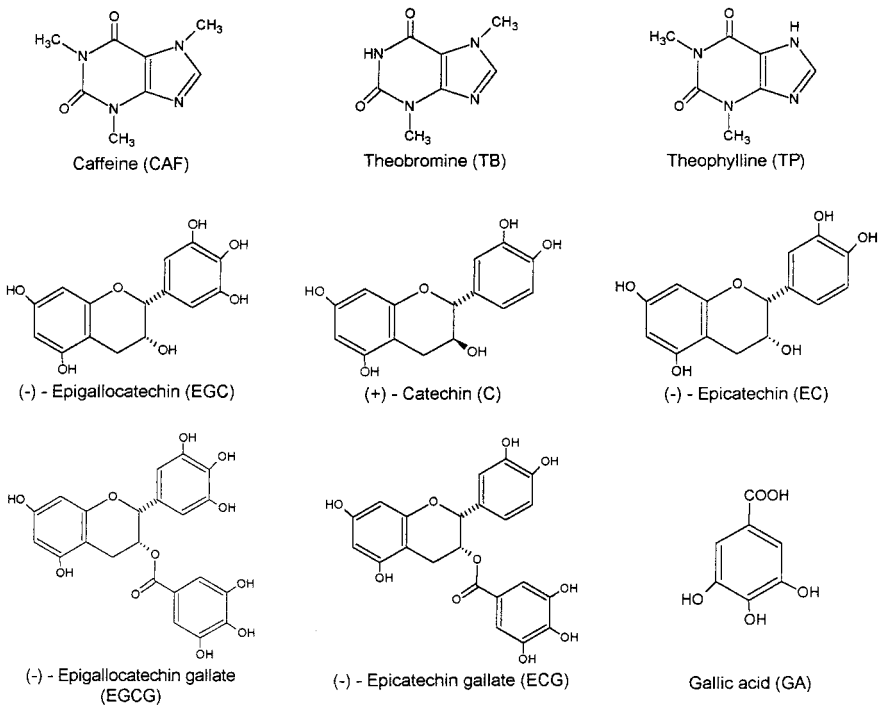

Figure 1. Structures of gallic acid, catechins, and tea alkaloids.

such as principal component analysis (PCA) and linear discriminant analysis (LDA), were applied for classification purposes.

MATERIALS AND METHODS

Apparatus. A Merck Hitachi (Darmstadt, Germany) HPLC chromatograph, equipped with a L-7100 pump, a Rheodyne (Cotati, CA) 7725i injection valve with a 20-μL sample loop, and a diode array detector L-7455 operating at 275 nm was used for the determinations. All the modules were controlled by a personal computer equipped with a Merck Hitachi D-7000 interface and HPLC System Manager software (Merck Hitachi). A 25 cm $\times$ 4 mm LichroCART RP-18 5-μm column (Merck) was used for the separation.

Reagents and Standard Solutions. Formic acid was obtained from Panreac, Barcelona, Spain. Acetonitrile (Romil, Cambridge, UK) was of HPLC grade. Milli-Q (Millipore, Bedford, MA) treated water with a resistivity higher than 18 MΩ cm was used throughout. Gallic acid, (+)-catechin, and (−)-epicatechin were purchased from Fluka (Buchs, Switzerland). Theobromine, theophillyne, (−)-epigallocatechin, (−)-epicatechin gallate, and (−)-epigallocatechin gallate were acquired from Sigma (Steinheim, Germany). These reagents were stored at −20 °C. Caffeine was obtained from Merck. Stock standard solutions (200 μg mL^{-1}) were prepared in acetonitrile and stored at 4 °C. Working standard solutions were prepared weekly from the stock solutions by dilution with acetonitrile.

Samples. A set of 45 commercial tea samples obtained from herbalists and specialized markets was selected for the analysis. Of the set of 45, 13 samples were nonfermented teas (green) and 32 were fermented teas (black and red). In all cases, the countries of origin were known, being China, Japan, Sri Lanka, Kenya, and India. In the case of teas from China and India, it was also known from which region of the country the teas were cultivated. In relation to the technology used in the manufacture of black teas, it was not possible to know which were CTC or orthodox. **Table 1** shows a short description of the samples and the corresponding code assigned to each one.

Sample Preparation. The tea samples were extracted according to the following procedure: 0.5 g of tea sample was extracted with 100 mL of acetonitrile/water (60:40, v/v) at room temperature for 1 h with constant stirring. The extract was filtered and diluted to volume in a 100-mL calibrated flask. Aliquots of 10 mL of this solution were transferred to a 25-mL volumetric flask rising with ultrapure water. Portions of this solution were filtered through a disposable 0.45-μm filter unit into a vial and injected into the HPLC system. This extraction procedure was carried out at room temperature to prevent possible degradation of the catechins (*22*). To assess the rudegeness of the extraction procedure, some samples were extracted within a weight range of 0.35−0.75 g. No significant differences were found in the determined contents of catechins and xanthines, regardless of the sample size taken from the extraction. As the results have been expressed on a dry basis, the moisture of the tea samples was determined before the analysis.

HPLC Method. The analytical determinations of gallic acid, theobromine, theophylline, catechin, caffeine, epicatechin, epigallocatechin, epicatechin gallate, and epigallocatechin gallate were carried out by means of high-performance liquid chromatography using a two-solvent gradient elution (*11*). The solvents compositions were (A) water/acetonitrile/formic acid (94.7:4.3:1 v/v) and (B) water/acetonitrile/formic acid (49.5:49.5:1 v/v). The mobile phase composition started at 90% solvent A and 10% solvent B, was increased linearly to 30% solvent B in 10 min, followed by a linear increase of solvent B to 80% in 5 min, and the final conditions being held for an additional 3 min. The flow rate was always 1 mL min^{-1}. To check the performance of the method the selectivity, linearity, detection and quantification limits, accuracy, and precision were evaluated. Resolutions of the peaks were calculated, and were higher than 1.5 in all the cases. The limit of detection (LOD) and limit of quantification (LOQ) were calculated as the concentration giving a signal equal to 3 and 10 times, respectively, of the signal/noise ratio (*23*). The accuracy of the method was evaluated from recovery assays, preparing spiked tea samples in triplicate at several levels of concentration. The average recoveries were calculated according to Cuadros et al. (*24*). **Table 2** summarizes the linear range, LOD, LOQ, and recovery values. The precision of the method was evaluated by carrying out eight replicate analysis of a standard solution on different days. The obtained relative standard deviation was always $<$2% for all the compounds.

Data Analysis. To carry out the chemometric calculations, two data matrices were built, one with 13 rows (nonfermented tea samples) and the other with 32 rows (nonfermented tea samples), and both with 9 columns (the analyzed chemical descriptors). Different pattern recognition techniques were applied to the data set. Principal component analysis (PCA) (*25*) was used to visualize the data trends. Linear discriminant analysis (LDA) (*26*) was applied as supervised learning

Table 1. Analyzed Tea Samples

code	origin	class[a]
1F	China (Fujian)	black
2F	China (Fujian)	black
3F	China (Fujian)	black
4F	China (Fujian)	black
5F	China (Lapsang Souchong, Fujian)	black
6F	China (Lapsang Souchong Smoky, Fujian)	black
1Chg	China (Gunpowder)	green
2Chg	China	green
3Chg	China	green
4Chg	China (Gunpowder)	green
5Chg	China (Jasmine)	green
6Chg	China	green
7Chg	China (Gunpowder)	green
1J	Japan	green
2J	Japan (Sencha)	green
3J	Japan	green
4J	Japan	green
5J	Japan	green
6J	Japan	green
1K	Kenya	black
2K	Kenya	black
3K	Kenya	black
4K	Kenya	black
5K	Kenya	black
6K	Kenya	black
1A	India (Assam)	black
2A	India (Assam)	black
3A	India (Assam)	black
4A	India (Assam)	black
5A	India (Assam)	black
1D	India (Darjeeling)	black
2D	India (Darjeeling)	black
3D	India (Darjeeling)	black
4D	India (Darjeeling)	black
1C	Sri Lanka (Ceylon)	black
2C	Sri Lanka (Ceylon)	black
3C	Sri Lanka (Ceylon)	black
4C	Sri Lanka (Ceylon)	black
5C	Sri Lanka (Ceylon)	black
6C	Sri Lanka (Ceylon)	black
1Y	China (Pu-erh, Yunnan)	red
2Y	China (Pu-erh, Yunnan)	red
3Y	China (Pu-erh, Yunnan)	red
4Y	China (Pu-erh, Yunnan)	red
5Y	China (Pu-erh, Yunnan)	red

[a] Nonfermented teas, green; fermented teas, black and red.

Table 2. Calibration and Sensitivity Parameters for the Determination of Gallic Acid, Catechins, and Methylxanthines

compound[a]	linear range (μg mL^{-1})	LOD[b] (ng mL^{-1})	LOQ[c] (ng mL^{-1})	recovery (%)
GA	0.5–45	4.8	16	98.2
TB	0.4–30	5.1	17	97.1
TP	0.1–18	4.9	16.3	92.6
EGC	1.5–90	90	300	99.7
C	1–10	30	100	95.3
CAF	1–120	4.8	16	100.3
EC	0.5–50	20	66.6	98.7
EGCG	0.5–0.84	16	53.3	101.2
ECG	0.5–72	4.3	14.3	99.4

[a] GA, gallic acid; TB, theobromine; TP, theophylline; EGC, epigallocatechin; C, catechin; CAF, caffeine; EC, epicatechin; EGCG, epigallocatechin gallate; ECG, epicatechin gallate. [b] LOD, limit of detection. [c] LOQ, limit of quantification.

method to find classification rules. The statistical package STATISTICA 99 from StatSoft (*27*) was used for pattern recognition calculations.

RESULTS AND DISCUSSION

Gallic Acid, Theophylline, Theobromine, Caffeine and Catechins Contents. The contents of the mentioned analytes present in the tea samples were determined using the HPLC method described in the previous section. **Figure 2** shows the chromatogram of standards obtained under the chromatographic

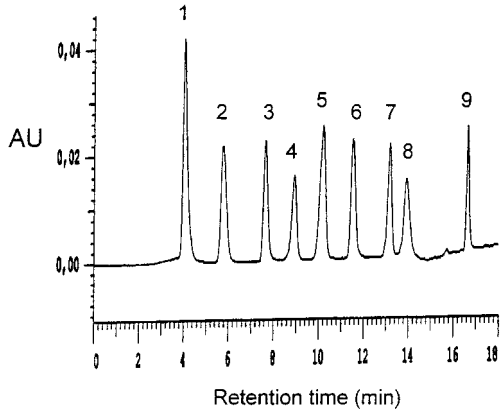

Figure 2. Chromatogram of a standard solution: (1) gallic acid (10 μg mL^{-1}), (2) theobromine (5 μg mL^{-1}), (3) theophylline (5 μg mL^{-1}), (4) epigallocatechin (60 μg mL^{-1}), (5) catechin (30 μg mL^{-1}), (6) caffeine (5 μg mL^{-1}), (7) epicatechin (20 μg mL^{-1}), (8) epigallocatechin gallate (10 μg mL^{-1}), and (9) epicatechin gallate (5 μg mL^{-1}).

Table 3. Contents (% w/w, dry basis) of Gallic Acid, Caffeine, Theophylline, Theobromine, and Catechins in the Analyzed Tea Samples[a]

sample	GA	TB	TP	EGC	C	CAF	EC	EGCG	ECG
1F	0.319	0.103	0.004	Nd[b]	0.195	3.907	0.346	1.623	0.757
2F	0.255	0.049	Nd	0.406	0.234	2.953	0.424	1.171	0.752
3F	0.226	0.047	Nd	Nd	0.175	3.409	0.198	0.532	0.542
4F	0.139	0.075	Nd	Nd	0.110	2.994	0.123	0.338	0.374
5F	0.268	0.084	Nd	Nd	0.073	3.324	0.123	0.606	0.274
6F	0.287	0.055	Nd	0.162	0.074	3.615	0.056	0.339	0.186
1Chg	0.129	0.066	Nd	3.368	0.270	2.631	0.874	7.386	1.361
2Chg	0.068	0.059	Nd	4.124	0.093	2.215	0.979	7.563	1.390
3Chg	0.164	0.203	Nd	3.843	0.552	3.422	2.161	8.729	3.556
4Chg	0.103	0.067	Nd	3.219	0.287	3.301	1.059	7.145	1.469
5Chg	0.168	0.123	Nd	3.004	0.564	3.863	1.413	9.396	2.525
6Chg	0.086	0.177	Nd	4.283	0.382	3.316	1.974	10.196	3.287
7Chg	0.107	0.046	Nd	2.291	0.130	2.596	0.735	4.943	0.915
1J	0.075	0.043	Nd	4.277	0.161	2.624	0.967	6.353	1.013
2J	0.038	0.002	Nd	3.840	0.188	2.194	0.812	7.533	1.264
3J	0.032	0.036	Nd	3.683	0.217	2.107	0.849	7.188	1.381
4J	0.070	0.086	Nd	4.300	0.356	2.896	0.963	7.354	1.198
5J	0.004	0.006	Nd	6.522	0.078	1.736	1.215	6.173	0.910
6J	0.018	0.018	Nd	4.659	0.093	1.468	1.193	5.693	1.115
1K	0.304	0.233	0.037	0.321	0.236	4.061	0.411	1.691	0.795
2K	0.253	0.196	0.040	0.259	0.178	3.688	0.187	1.133	0.580
3K	0.212	0.255	0.042	0.203	0.206	4.465	0.244	1.231	0.674
4K	0.303	0.226	0.044	0.287	0.213	4.165	0.311	1.584	0.806
5K	0.291	0.219	0.049	0.281	0.241	4.234	0.238	1.442	0.682
6K	0.193	0.244	0.041	0.178	0.189	4.423	0.217	1.069	0.512
1A	0.351	0.460	0.029	0.187	0.098	4.728	0.072	0.938	0.625
2A	0.190	0.366	0.022	Nd	0.070	4.774	0.026	0.583	0.377
3A	0.154	0.249	0.022	Nd	0.074	4.070	0.010	0.493	0.255
4A	0.179	0.335	0.030	Nd	0.106	4.614	0.063	0.834	0.467
5A	0.175	0.323	0.031	Nd	0.101	4.862	0.070	0.869	0.471
1D	0.235	0.147	0.027	0.679	0.242	4.212	0.378	6.156	1.722
2D	0.241	0.106	0.028	0.522	0.310	3.740	0.303	5.080	1.535
3D	0.226	0.191	0.028	1.077	0.466	4.228	0.506	8.233	2.434
4D	0.358	0.120	0.033	0.274	0.149	4.042	0.180	2.281	0.978
1C	0.293	0.177	Nd	0.895	0.165	3.722	0.580	1.515	0.650
2C	0.266	0.163	0.036	Nd	0.176	3.621	0.390	0.841	0.850
3C	0.277	0.193	0.033	0.506	0.213	3.696	0.516	1.627	0.937
4C	0.224	0.221	0.026	Nd	0.139	4.027	0.162	0.522	0.609
5C	0.198	0.103	0.039	2.955	0.256	3.360	1.735	4.813	1.811
6C	0.194	0.173	0.043	0.318	0.181	3.406	0.422	1.914	1.105
1Y	0.104	0.115	0.041	0.147	0.043	3.640	0.087	0.062	0.059
2Y	0.496	0.130	0.036	0.360	0.068	3.862	0.233	0.169	0.104
3Y	2.537	0.299	0.038	0.786	0.317	3.675	0.853	0.288	0.282
4Y	1.697	0.224	0.043	0.992	0.266	3.809	0.970	0.377	0.226
5Y	1.026	0.237	0.046	1.068	0.103	2.415	0.531	0.297	0.167

[a] GA, gallic acid; TB, theobromine; TP, theophylline; EGC, epigallocatechin; C, catechin; CAF, caffeine; EC, epicatechin; EGCG, epigallocatechin gallate; ECG, epicatechin gallate. [b] Not detected.

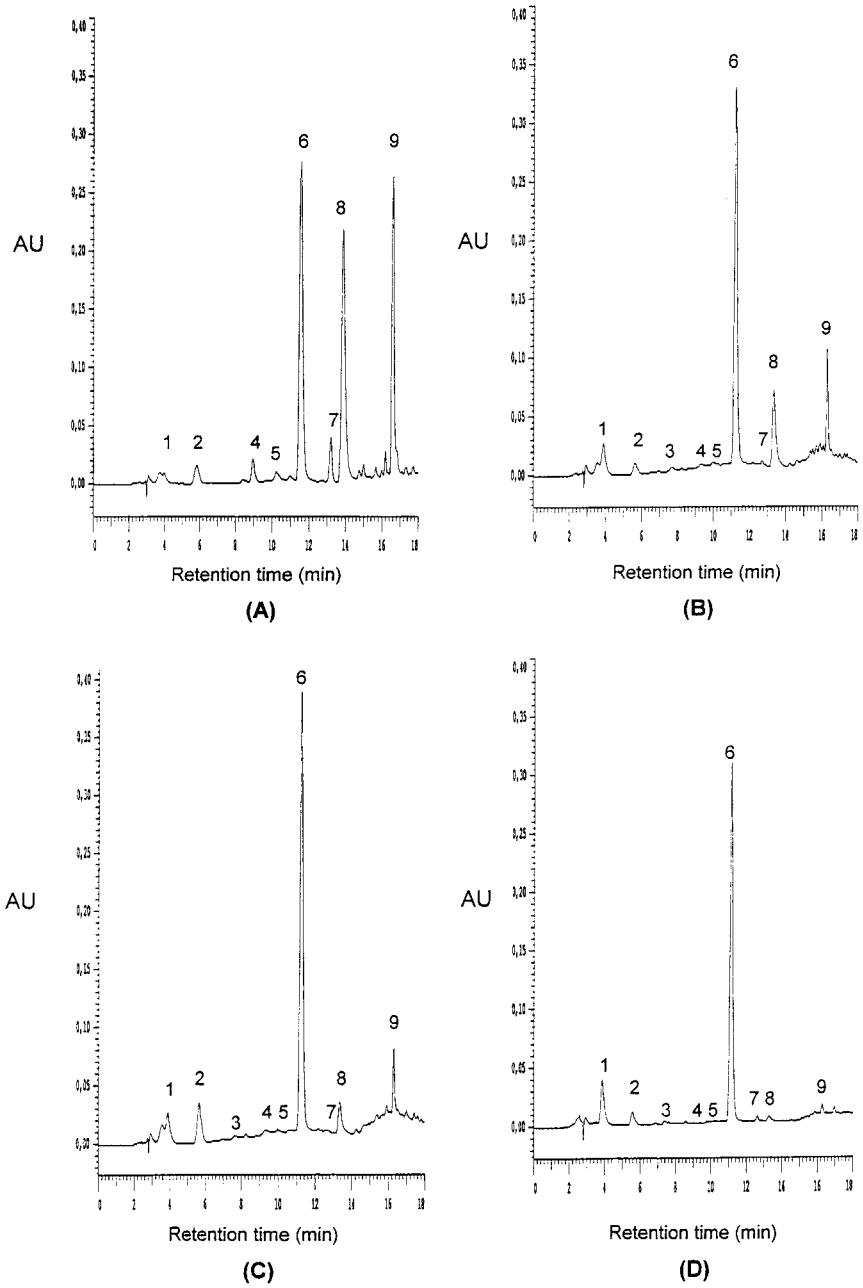

Figure 3. Chromatograms of fermented and nonfermented tea samples of different geographical origins: (A) China (nonfermented); (B) India (Darjeeling, fermented); (C) India (Assam, fermented); (D) China (Yunnan, fermented). Peaks are (1) gallic acid, (2) theobromine, (3) theophylline, (4) epigallocatechin, (5) catechin, (6) caffeine, (7) epicatechin, (8) epigallocatechin gallate, and (9) epicatechin gallate.

conditions described above. The analyzed compounds were identified in the tea samples by comparing their retention times with those of standard solutions. Under the operating conditions, the retention times (in minutes) for the studied compounds were as follows: 3.9 (GA), 5.5 (TB), 7.3 (TP), 8.5 (EGC), 9.7 (C), 11.0 (CAF), 12.5 (EC), 13.3 (EGCG), and 16.3 (ECG). Furthermore, standard additions of each compound were carried out to confirm the presence of the compounds in the samples. In **Figure 3**, chromatograms of fermented and non-fermented tea samples from different geographical origins are presented. Clear differences among the profiles can be observed. Nonfermented teas present higher peaks for all the catechins, especially EC, EGCG, and ECG. The chromatograms corresponding to fermented teas always show lower peaks of catechins. It is also remarkable that peaks corresponding to gallic acid are always higher in fermented teas, specially in the case of China teas from Yunnan. This can be explained by the fact that the amount of gallic acid increases during the fermentation process owing to its liberation from catechin gallates (*8*). On the other hand, the peak corresponding to caffeine is the highest in all cases. The results obtained after analyzing the tea samples, expressed in % w/w, dry basis, are depicted in **Table 3**. Gallic acid presents values that range between 0.004 and 2.537%, and the lower percentages are obtained for the nonfermented teas. Yunnan teas present the highest values of gallic acid. The amount of theobromine ranges between 0.002 and 0.460%. Theophylline cannot be detected in nonfermented teas and in China teas from Fujian (except sample 1F). For the rest of the samples, the theophylline content is always lower than 0.049%. Caffeine is present in higher amounts in the case of fermented teas, showing values between 2.415 and 4.862%, whereas nonfermented teas show caffeine levels ranging between 1.468 and 3.863%. The amounts of catechins are always higher for nonfermented teas. EGCG and EGC are the major catechins present with average

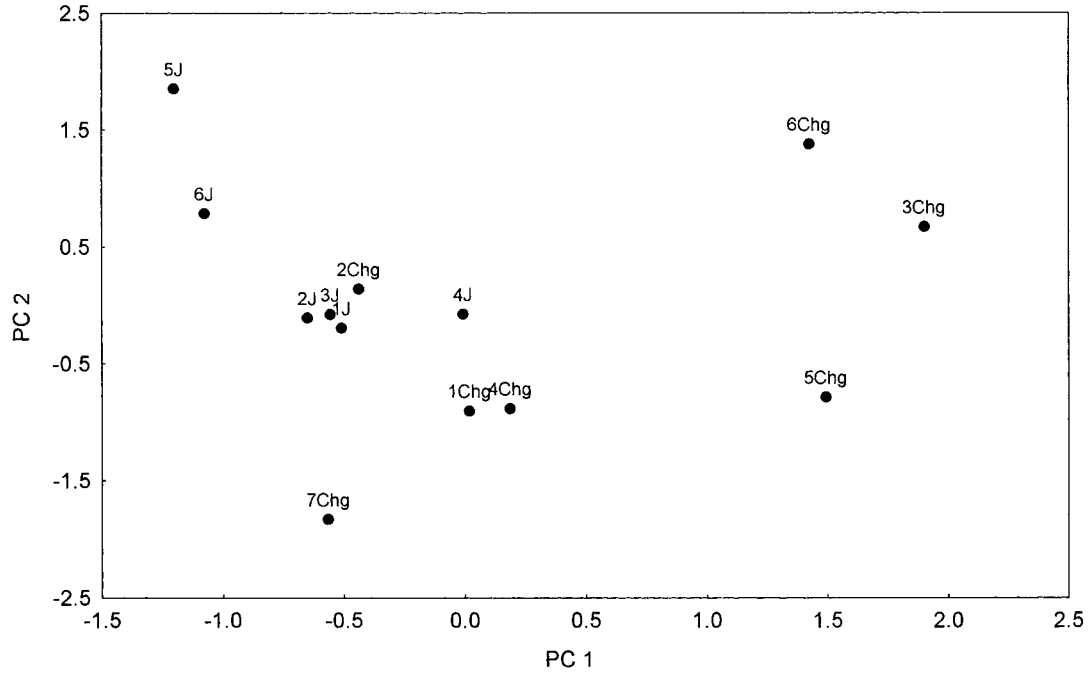

Figure 4. Scores plot of the nonfermented tea samples for the first PCs.

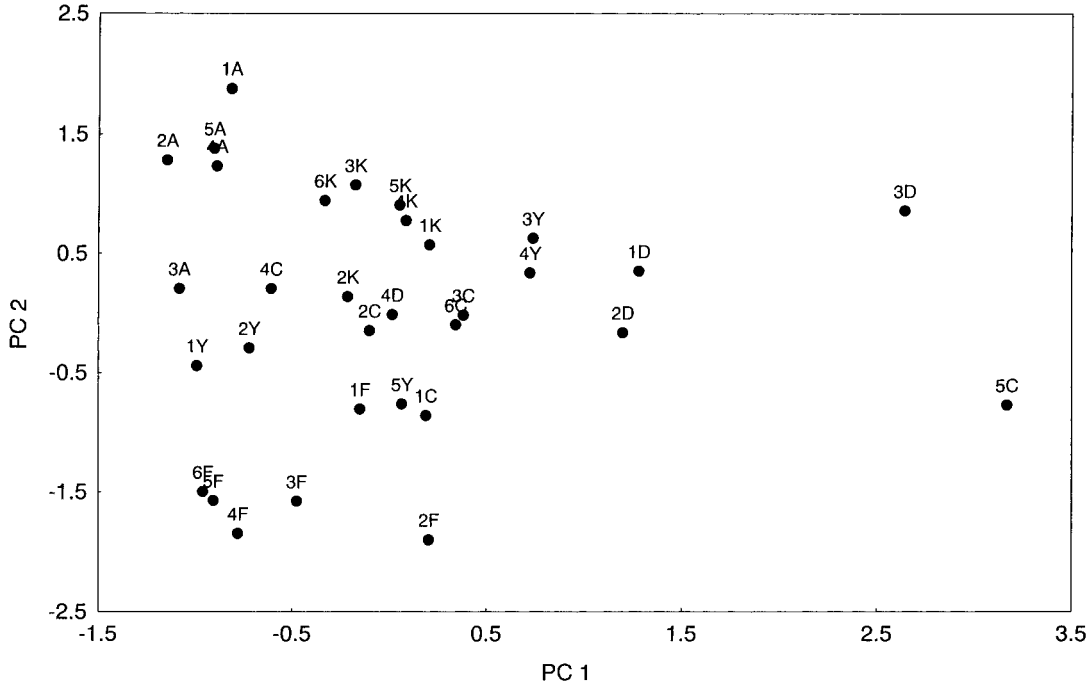

Figure 5. Scores plot of the fermented tea samples for the first PCs.

contents of 7.358% and 3.955%, respectively. ECG presents values ranging between 3.556 and 0.910%. For fermented teas, EGCG and ECG are the catechins present in major percentages, with average contents of 1.583 and 0.706%, respectively, and these values are less than those found in nonfermented teas. EGC cannot be detected in the majority of Assam and Fujian teas or in some Ceylon teas (2C and 3C).

Geographical Classification of Teas. For chemometric calculations, the tea samples were divided into two groups, forming two different data sets. The nonfermented teas, i.e., teas from Japan and China (13 green teas), compose one of them. The fermented teas (27 black and 5 red teas) constitute the other data set (32 samples). Pattern recognition procedures were applied to these data sets, trying to classify the tea samples according to their geographical origin.

Principal Component Analysis. This technique was applied in order to visualize the data trends and provides a first evaluation of the discriminant efficiency of the selected features. PCA is based on the derivation of linear combinations of the measured descriptors to produce new variables called principal components (PCs) that are uncorrelated. PCs are obtained sequentially: the first PC (PC1) accounts for the largest portion of explainable variability in the measured data, the second one (PC2) accounts for the next largest portion of explainable data variability, and so forth. In other words, PCA attempts to condense the information (variability) of measured data ex-

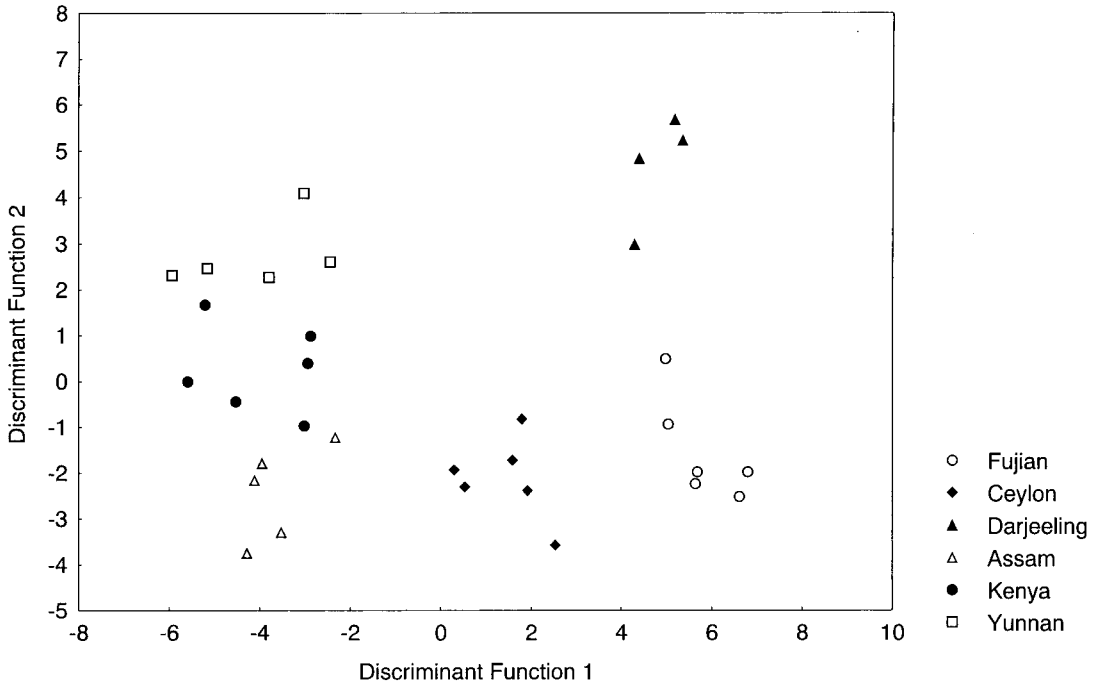

Figure 6. Plot of the fermented tea samples in two of the discriminant functions.

plained by the first PCs. Hence, data plots using PC1 and PC2 as variables (scores plot) enable us to visualize the data trends of the data matrix with a lesser dimensionality (*28*, *29*). After applying this method to the data set formed by the nonfermented teas, two principal components were extracted, explaining up to 71.44% of the total variance. The equations of these two PCs are as follows:

$$\text{PC1} = -0.35\ \text{EGC} + 0.94\ \text{C} + 0.89\ \text{CAF} + 0.78\ \text{EC} + 0.84\ \text{EGCG} + 0.93\ \text{ECG} + 0.82\ \text{GA} + 0.96\ \text{TB}$$

$$\text{PC2} = 0.88\ \text{EGC} - 0.07\ \text{C} - 0.31\ \text{CAF} + 0.55\ \text{EC} + 0.60\ \text{EGCG} + 0.28\ \text{ECG} - 0.48\ \text{GA} + 0.13\ \text{TB}$$

Figure 4 shows the corresponding scores plot obtained from eigenanalysis of the covariance matrix of autoscaled data. At a glance, certain separation between Chinese and Japanese teas can be observed. Teas from Japan are located at the left side of the plot, at negative values of PC1. Most of the Chinese teas are situated at positive values of PC1, except samples 7Chg and 2Chg that appear quite close to the Japanese teas cluster. Thus, a certain grouping of teas according to their origin can be noticed. After examination of the variable loadings of PC1, TB, C, ECG, CAF, and EGCG are the descriptors with more contribution to this PC.

When PCA was applied to the data set formed by the fermented teas, the two first PCs explained 65.4% of the variance. The equations of these PCs are as follows:

$$\text{PC1} = 0.18\ \text{GA} - 0.20\ \text{TB} + 0.24\ \text{TP} + 0.83\ \text{EGC} + 0.84\ \text{C} - 0.17\ \text{CAF} + 0.81\ \text{EC} + 0.81\ \text{EGCG} + 0.83\ \text{ECG}$$

$$\text{PC2} = 0.09\ \text{GA} + 0.87\ \text{TB} + 0.64\ \text{TP} - 0.10\ \text{EGC} + 0.16\ \text{C} + 0.84\ \text{CAF} - 0.13\ \text{EC} + 0.13\ \text{EGCG} + 0.12\ \text{ECG}$$

The variables with higher contribution to PC1 are C, EGC, EC, ECG, and EGCG. **Figure 5** shows the corresponding scores plot. Though there is no clear separation between the samples, some trends can be observed. The Chinese teas from Fujian appear at the left lower side of the plot, whereas the Kenyan teas are grouped together in the center of the plot. Assam teas are situated at the left upper side. Darjeeling, Ceylon, and Yunnan teas do not form defined clusters.

Linear Discriminant Analysis. LDA differs from data reduction methods such as PCA in that it is concerned with determining the so-called discriminant functions as linear combinations of the chemical descriptors which best separate the classes according to minimization of the ratio of within-class and between-class sum of squares. The number of discriminant functions (t) is the minimum value between the number of classes less one ($k - 1$) and the number of descriptors, p ($t = \min(k - 1, p)$) (*30*). An a priori knowledge of the number of classes and the class membership of each sample in the data matrix is assumed. Once the discriminant functions (DF) have been calculated, the classes are separated by hyperplanes into subspaces within the space of the discriminant functions, and the samples are classified according to falling into one of the class subspaces (classification rule). First, we considered the data matrix corresponding to the nonfermented teas, thus there are two classes in this case (i.e., Chinese green teas and teas from Japan). The obtained discriminant function was

$$\text{DF1} = 1.82\ \text{GA} + 0.38\ \text{TB} - 0.94\ \text{EGC} - 3.03\ \text{C} + 0.14\ \text{CAF} + 1.90\ \text{EC} + 1.65\ \text{EGCG} - 1.37\ \text{ECG}$$

Only one sample (1J) was misclassified according to the a posteriori probabilities which represent a recognition ability of 92.3%. The leave-one-out method (*31*) was used as a validation procedure and also to evaluate the classification performance, obtaining a prediction ability of 92.3%. On the other hand, when applying LDA to the data matrix composed by the fermented teas, six classes were considered according to the different geographical origins of the studied tea samples, i.e., Fujian, Ceylon, Darjeeling, Assam, Kenya, and Yunnan. Thus, five discriminant functions can be calculated. The theophylline content in these samples was not considered for calculations because it was not detected in several of the studied samples.

The corresponding equations for DF1 and DF2 are as follows:

$$DF1 = 1.60\,GA - 1.11\,TB - 1.08\,TP - 0.44\,EGC - 1.54\,C - 0.11\,CAF - 0.84\,EC - 1.40\,EGCG + 3.82\,ECG$$

$$DF2 = 1.62\,GA - 0.89\,TB + 0.71\,TP + 1.16\,EGC + 0.04\,C + 0.15\,CAF - 2.34\,EC + 1.95\,EGCG - 1.05\,ECG$$

If we plot the samples in the plane defined by DF1 and DF2, we obtain a distribution that is shown in **Figure 6**. It can be noticed that samples of the same geographical origin are grouped in the same clusters and are well separated from the other clusters composed of teas from different geographical areas. Looking at the a posteriori probabilities, a 100% of recognition ability was obtained. After applying the leave-one-out method, only four samples were misclassified, representing 87.5% of prediction ability.

Thus, it can be concluded that catechins, gallic acid, and methyl-xanthines are very adequate descriptors to differentiate tea samples from different geographical origins as well as to establish classification rules for these purposes. Nevertheless, further research using a higher number of samples, covering aspects such as the technology used in the manufacturing and seasonal variation, would be necessary to confirm the connections between the contents of catechins and xanthines with the geographical origin.

LITERATURE CITED

(1) Grahan, H. N. The polyphenols of tea-biochemistry and significance − a review. In *Xve Journées Internationales Group Polyphenols*; DTA: Lisbon, Portugal, 1992; Vol. 2, pp 32−43.

(2) Lin, J. K.; Lin, C. L.; Liang, Y. C.; Lin-Shiau, S. Y.; Juan, I. M. Survey of catechins, gallic acid, and methylxanthines in green, oolong, pu-erh, and black teas. *J. Agric. Food Chem.* **1998**, *46*, 3635−3642.

(3) Lin, Y. L.; Juan, I. M.; Chem, Y. L.; Liang, Y. C.; Lin, J. K. Composition of polyphenols in fresh tea leaves and associations of their oxygen-radical-absorbing capacity with antiproliferative actions in fibroblast cells. *J. Agric. Food Chem.* **1996**, *44*, 1387−1394.

(4) Ho, C. T.; Chen, Q.; Shi-Zhang, K. Q.; Rosen, R. T. Antioxidative effect of polyphenol extract prepared from various Chinese teas. *Prev. Med.* **1992**, *21*, 520−525.

(5) Henry, J. P.; Stephens-Larson, P. Reduction of chronic psychosocial hypertension in mice by decaffeinated tea. *Hypertension* **1984**, *6*, 437−444.

(6) Chen, J.; Han, C. The protective effect of tea on cancer: Human evidence. In *Phytochemicals as Bioactive Agents*; Bidlach, W. R., Omaye, S. T., Meshin, M. S., Topham, D. K. W., Eds.; Technomic: Lancaster, PA, 2000; pp 131−150.

(7) Lin, Y. L.; Lin, J. K. (−)-Epigallocatechin-3-gallate blocks the induction of nitric oxide synthase by down-regulating lipopolysaccharide-induced activity of transcription factor NF-kB. *Mol. Pharmacol.* **1997**, *52*, 465−472.

(8) Xie, M.; Von Bohlen, A.; Klockenkämper, R.; Jian, X.; Günther, K. Multielement analysis of Chinese tea (*Camellia sinensis*) by total-reflection X-ray fluorescence. *Z. Lebensm.-Unters-Forsch.* **1998**, *207*, 31−38.

(9) Dalluge, J. J.; Nelson, B. C.; Thomas, J. B.; Sander, L. C. Selection of column and gradient elution system for separation of catechins in green tea using high performance liquid chromatography and a modified digestion procedure. *J. Chromatogr. A* **1998**, *783*, 265−274.

(10) Finger, A.; Kuhr, S.; Engelhardt, U. H. Chromatography of tea constituents. *J. Chromatogr.* **1992**, *624*, 293−315.

(11) Fernández, P. L.; Martín, M. J.; González, A. G.; Pablos, F. HPLC determination of catechins and caffeine in tea. Differentiation of green, black and instant teas. *Analyst* 2000, *125*, 421−425.

(12) Horie, H.; Mukai, T.; Kohata, K. Simultaneous determination of qualitative important components in green tea infusions using capillary electrophoresis. *J. Chromatogr. A* **1997**, *758*, 332−335.

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

(14) Arce, L.; Ríos, A.; Valcárcel, M. Determination of anticarcinogenic polyphenols present in green tea using capillary electrophoresis coupled to a flow injection system. *J. Chromatogr. A* **1998**, *827*, 113−120.

(15) Wang, H.; Helliwell, K.; You, X. Isocratic elution system for the determination of catechins, caffeine and gallic acid in green tea using HPLC. *Food Chem.* **2000**, *68*, 115−121.

(16) Goto, T.; Yoshida, Y.; Kiso, M.; Nagashima, H. Simultaneous analysis of individual catechins and caffeine in green tea. *J. Chromatogr. A* **1996**, *749*, 295−299.

(17) Kuhr, S.; Engelhardt, U. H. Determination of flavanols, theogallin, gallic acid and caffeine in tea using HPLC. *Z. Lebensm.-Unters.-Forsch.* **1991**, *192*, 526−529.

(18) Bronner, W. E.; Beecher, G. R. Method for determining the content of catechins in tea infusions by high-performance liquid chromatography. *J. Chromatogr. A* **1998**, *805*, 137−142.

(19) Shao, W.; Powell, C.; Clifford, M. N. The analysis by HPLC of green, black and pu'er teas produced in Yunnan. *J. Sci. Food Agric.* **1995**, *69*, 535−540.

(20) Marcos, A.; Fisher, A.; Rea, G.; Hill, S. Preliminary study using trace element concentrations and a chemometrics approach to determine the geographical origin of tea. *J. Anal. At. Spectrom.* **1998**, *13*, 521−525.

(21) Fernández-Cáceres, P. L.; Martín, M. J.; Pablos, F.; González, A. G. Differentiation of tea (*Camellia sinensis*) varieties and their geographical origin according to their metal content. *J. Agric. Food Chem.* **2001**, *49*, 4775−4779.

(22) Suematsu, S.; Hisanobu, Y.; Saigo, H.; Matsuda, R.; Komatsu, Y. A new extraction procedure for determination of caffeine and catechins in green tea. *Nippon Shokunin Kagaku Kogaku Kaishi* **1995**, *42*, 419−424.

(23) Miller, J. C.; Miller, J. N. Errors in instrumental analysis. In *Statistics for Analytical Chemistry*; Ellis Horwood: Chichester, U. K., 1988; pp 115−117.

(24) Cuadros, L.; García Campaña, A. M.; Alés, F.; Jiménez, C.; Román Ceba, M. *J.* Validation of an analytical instrumental method by Standard Addition Methodology. *AOAC Int.* 1995, *78*, 471−476.

(25) Chatfield, C., Collins, A. J.; Eds. *Introduction to Multivarite Analysis*; Chapman & Hall: London, 1980; pp 57−81, Principal Component Analysis.

(26) Coomans, D.; Massart, D. L.; Kaufman, L. Optimization by statistical linear discriminant analysis in analytical chamistry. *Anal. Chim. Acta* **1979**, *112*, 97−122.

(27) STATISTICA for Windows (Computer Program Manual); StatSoft, Inc.: Tulsa, OK, 1999; (http://www.statsoft.com).

(28) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37−52.

(29) Gardiner, W. P. *Statistical Analysis Methods for Chemists*; Royal Society of Chemistry: Cambridge, U. K., 1997.

(30) González-Arjona, D.; González, A. G. Adaptation of linear discriminant analysis to second level pattern recognition classification. *Anal. Chim. Acta*. **1998**, *363*, 89−95.

(31) Henrion, R.; Henrion, G. Überwachte Klassifikation. In *Multivariate Datenanalysen*; Springer-Verlag: Berlin, Germany, 1995; pp 71−73.

Received for review November 1, 2001. Revised manuscript received January 22, 2002. Accepted January 22, 2002.

JF0114435