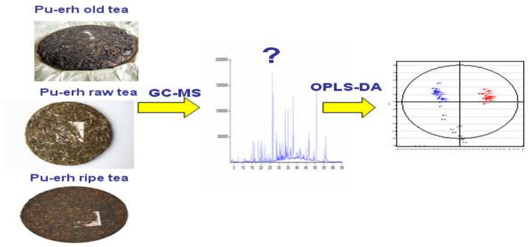
RSC Advances

ROYAL SOCIETY
OF CHEMISTRY

www.rsc.org/advances



Graphical abstracts

Application of gas chromatography-mass spectrometer and chemometrics methods for assessing volatile profiles of Pu-erh tea with different processing ways and ageing year

Shidong Lv [1,2] . Yuanshuang Wu[1] . Jifu Wei[3]. Ming Lian[1] . Chen Wang[1]. Xuemei Gao[1] . Qingxiong Meng[1*]

[1] Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming 650500, Yunnan, People's Republic of China

[2]Kunming Grain & Oil and Feed Product Quality Inspection Center, Kunming 650118, Yunnan, People's Republic of China

[3] Research Division of Clinical Pharmacology, the first affiliated hospital, Nanjing Medical University, Nanjing 210029, People's Republic of China

* Corresponding author: Qingxiong Meng

Tel: +86-871-65920541, Fax: +86-871-65920570, E-mail: qxmeng@scbg.ac.cn

Abstract

Volatile changes and the post-fermentation ageing process of tea remain largely unknown. Additionally, the understanding of the ageing and storage processes of tea mostly rely on sensory experience and lack the support of scientific and accurate data. In this paper, a development method based on the head-space solid phase microextraction/gas chromatography-mass spectrometer (HS-SPME/ GC-MS) combined with multivariate statistical methods was used to assess volatile profiles in different type Pu-erh teas, including raw, ripe and aged Pu-erh teas. A total of 122 aroma components were identified in 57 Pu-erh teas. Differences in the manufacturing type and storage years of Pu-erh teas resulted in different compositions and contents of volatile components. The characteristic volatiles in aged teas were hexadecanoic acid, dihydroactinidiolide, caffeine, linalool, 6,10,14-trimethyl-2-pentadecanone, β-ionone, cedrol, and phytol; the characteristic volatiles in raw teas were linalool, tridecane, caffeine, dihydroactinidiolide, β-ionone, 6,10,14-trimethyl-2-pentadecanone, dodecane, etc.; and the characteristic volatiles in ripe teas were 1,2,3-trimethoxybenzene, hexadecanoic acid, 1,2,4-trimethoxybenzene, dihydroactinidiolide, 6,10,14-trimethyl-2-pentadecanon, caffeine, and 1,2,3-trimethoxy-5-methyl-benzene. Through principal component analysis (PCA), clustering analysis (CA) and orthogonal projection to latent structures discriminate analysis (OPLSDA), three different kinds of Pu-erh teas were classified successfully. Additionally, the aged Pu-erh teas showed similar volatile constituents as the ripe teas. This study suggested that HS-SPME/GCMS combined with chemometrics methods is accurate, sensitive, fast, universal and ideal for rapid routine analysis and discrimination of Pu-erh teas with different processing technologies and storage times.

Keywords: *Pu-erh tea; manufacturing type; volatile component; aged tea; HS-SPME; GC-MS; chemometrics methods*

Introduction

Pu-erh tea is made from the tea leaves of Yunnan big leaves species in a certain area of the Yunnan province [1]. Pu-erh tea is usually processed by two different methods; raw Pu-erh tea (also known as Pu-erh green tea) is traditionally produced by pressing the sun-dried green tea leaves. Ripe Pu-erh tea is produced by a pile fermentation process in a hot and humid environment with microorganisms for approximately 60 days before it is pressed. In addition, Pu-erh raw tea can have a similar flavour characteristic as Pu-erh ripe tea after a long ageing time (generally longer than 10 years) and is usually called Pu-erh aged tea or dry-stored tea [2, 3]. In recent years, more and more people like Pu-erh tea because of its unique flavour and its health efficacy; thus, Pu-erh tea has sold well and obtained a good reputation both in China and abroad [4].

The ageing of tea is common, and drinking aged tea seems likely to become a trend in some countries. Many teas, such as Anji white tea, Fuzhuan tea, High Mountain Oolong tea, Sichuan dark tea, and Pu-erh tea, have a better quality (taste and flavour) and better health efficacy if they are treated with long-term ageing. Pu-erh tea has been recognized since the Tang dynasty (AD 618-906) in China, and the taste has been observed to improve with longer preservation times [5]. Thus, Pu-erh aged tea is generally more expensive than newly produced Pu-erh tea. Tea consumers and merchants are often willing to pay higher prices to obtain older Pu-erh tea. With the increasing awareness of the importance of healthy living, the consumption demand of high quality Pu-erh aged tea products has been increasing significantly. To obtain higher profits, some manufacturers are misleading customers by labelling inferior or relatively new Pu-erh tea products as older Pu-erh tea. Sensory evaluation is currently a common method used to differentiate various teas that have undergone different processing methods [6]. However, this approach cannot always result in an objective judgement because it is easily influenced by physical and mental conditions [7]. Moreover, this approach is unable to reflect the chemical difference of various different ages of tea and usually lacks a comprehensive view of tea chemical compositions. Thus, developing a method for discriminating the natural ageing process and identifying the processing methods (raw tea and ripe tea) is urgently required not only for consumers to protect their interests but also for tea producers to apply quality control. Additionally, analysing the chemical composition changes with different productive years by using modern instrumental analysis methods can help us to understand the dynamic changes of chemical compounds in tea that is stored during the ageing process.

The rapid development of modern instrumental analysis technology, such as LC (liquid chromatogram), near infrared reflection (NIR), inductively coupled plasma mass spectrometry (ICP-MS), nuclear magnetic resonance (NMR), GC-MS, etc., make it possible to obtain more information (variables) of large samples (objects) in a relatively shorter time [8-10]. Consequently, making use of different multivariate statistical methods to maximize the extraction of useful information from large data sets is of great importance. These methods mainly include principal component analysis (PCA), cluster analysis (CA), SIMCA (soft independent modelling of class analogy), orthogonal partial least squares discriminate analysis (OPLS-DA), etc. [11-13]. Currently, instrumental analysis technologies combined with these chemometrics methods have been applied with increasing success in the analysis of samples of different tea types. For instance, using GC-MS and the fingerprinting technique combined with global volatile profiling, we have revealed that the green teas have a feature of producing area-dependent chemical components [14]. Lin *et al*. [15] successfully used HS-SPME combined with PCA, CA and linear discriminate analysis (LDA) to discriminate five Oolong tea (*Camellia sinensis*) varieties; Ye *et al*. [16] applied similar methods to distinguish between Hubei green teas and

Henan green teas in China. Other analysis technologies, such as NMR [17], ICP-MS [18], NIR [19], and LC-MS [20] combined with multivariate statistical methods also obtained satisfactory results when applied to identify geographic origins and processing types of various tea samples. Therefore, multivariate statistical methods combined with HS-SPME/GC-MS could be an effective and convenient tool for comprehensive analysis of different tea volatiles in order to assess tea quality and to investigate the behaviour of volatiles during long-term tea storage.

To our knowledge, the study of chemical composition associated with processing methods and production length of Pu-erh tea is very limited. Ku *et al*. [21] have used LC-MS and chemometrics methods to analyse the processing type and change of water-soluble components in Pu-erh tea with different post fermentation lengths, demonstrating that a chemometrics method was an effective tool for identifying processing types and the post fermentation length of Pu-erh tea. Xu *et al*. [22] used NIR spectroscopy combined with chemometrics methods to discriminate the type (raw tea and ripe tea) and predict the age of Chinese tuo tea. Nevertheless, there is no study on volatile profiling using GC-MS and multivariate statistical method analysis from different production years and processing types of Pu-erh tea, especially Pu-erh aged tea. Additionally, the study of ageing processes and the similarities and differences between Pu-erh aged tea and ripe tea are essentially unknown.

Our previous study [14, 23] used the HS-SPME/GC-MS methodology to investigate the fingerprint characteristic of Dayi Pu-erh ripe tea and Pu-erh green tea; however, it mainly involved the fingerprint similarity analysis method, such as the correlation coefficient of similarity and overlapping chromatographic peaks (ORP), and did not involve the OPLS-DA method; additionally, only a single processed type of Pu-erh tea was involved. In the present study, the GC-MS method combined with PCA, CA and OPLS-DA techniques was adopted to probe potential differences in aroma characteristics among three types of Pu-erh teas (raw, ripe and aged). Changes in the content of volatile components in relation to the ages and production process of Pu-erh teas were also explored. Our study might provide a reference point for a fast, simple method for identifying Pu-erh teas that have been produced with different ages and processing types.

Materials and methods

Materials

Different ages of Pu-erh raw, aged and ripe teas were collected from Jingmai Mountain in the Pu-erh district, located in the southwest Yunnan Province. The volatile compositions of tea are easily affected by climates, agricultural practices, and storage conditions; thus, tea samples coming from the same geographical area were collected to minimize the effects of these factors. Although the aged and ripe teas have a similar appearance, they undergo different fermentation methods. Ripe teas were rapidly pile-fermented under the action of microorganisms, whereas aged teas are piled for many years in a natural way without being processed by pile fermentation. A total of 57 Pu-erh tea samples were collected. Raw teas more than 10 years old were regarded as old-aged teas; these have obvious differences in appearance, flavour, and colour of tea infusion compared with fresh raw tea. To ensure maximum reliability of the production year of Pu-erh old tea samples, all of the samples that were provided must have a production logo (batch, production time, raw material, etc.) or collection records; at the same time, there must be two evaluation experts that further determined auxiliary factors in combination with their sensory characteristics (shape, liquor colour, taste, etc.). It is important to note that a time error may exist when the Pu-erh tea is aged more than 30 years. Because it was very difficult to find a reliable source of old teas, the number of old tea samples in this study was relatively small (only tens). The details of all these tea samples were listed in Table 1. Among 57-Pu-erh teas, "O" represented aged teas, "R" represented raw teas, and "P" represented ripe teas. Additionally, samples from the same year were presented with different codes. For example, O8 means the old aged tea from the year 2008, and O8-1 and O8-2 means two different samples from the year 2008. Because our research was performed in March 2014, the Pu-erh tea samples that were collected were produced before the year 2014.

HS-SPME method

The solid-phase extraction coatings (65 μm polydimethylsiloxane/divinylbenzene (PDMS/DVB)) were provided by Supelco (Bellefonte, PA, USA). The HS-SPME method was described in detail by a previous study [23]. The ground tea sample (2.0 g) was weighed and placed in a 20 mL sealed headspace vial; then the sample was infused with 5 mL of boiling distilled water. Then, the HS-SPME fibre was exposed to the sample headspace while the tea powder was continuously stirred (250 rpm) for 60 min at 80 °C. After extraction, the fibre coatings were removed from the headspace vial and were immediately inserted into the GC-MS splitless injector for absorbance (250 °C for 3.5 min) and further separation and identification.

GC-MS analysis

A 7890A GC-5975C MS system (Agilent Technologies, CA, USA) was used for separation and identification of volatile components of Pu-erh tea. The chromatographic column was an HP-5MS column (30 m × 0.25 mm × 0.25 μm film thickness), with the high purity helium acting as the gas carrier; the flow rate was controlled at 1 mL min^{-1}. The injector temperature was 250 °C and was equipped with a splitless injector. The temperature was programmed for 50 °C (held for 1 min) and increased to 210 °C at 3 °C min^{-1} (held for 3 min) and then was programmed for 210 to 230 °C for 15 °C min^{-1}. The MS ion source temperature was 230 °C, and electron energy was 70 eV. The scan range was 35–500 amu. The solvent delay time was 2.8 min.

Data processing

Chromatographic peaks were recognized by ChemStation and were identified by the NIST 08.L MS data library and the retention indices (RI) method [24-27]. The relative content of the chromatographic peaks were obtained by a peak area normalization method based on the total ion current. The RI of each compound was obtained by using a 1 μL n-alkane mixture (C8-C40; Sigma-Aldrich, USA) under the same GC-MS experimental conditions. The data matrix was transferred in to the SIMCA-P12 package (Umetrics, Umea, Sweden) where PCA, CA and OPLS-DA analyses were conducted. Duncan's multiple range tests were used to test the significance of the differences among different groups of Pu-erh tea samples using SPSS 17.0 software.

Results and discussion

Repeatability and stability test

The repeatability of the HS-SPME method was determined by analysing the same Pu-erh tea sample 6 times under the same experimental conditions. The relative standard deviation (% RSD) of the peak area for the volatile components ranged from 7.45 to 11.13%. With the same HS-SPME/GC–MS method, the sample stability was determined at 0, 4, 8, 16, 24, 48 h by using the same Pu-erh raw tea sample (R1). The RSD of the peak area for the volatile components ranged from 8.97 to 11.98%. The repeatability and stability test results indicated that the HS-SPME/GC-MS method was reliable and applicable for analysis of the volatile components of Pu-erh tea.

Analysis of volatile profiles of Pu-erh teas with different types

To investigate aroma characteristics of the tested Pu-erh tea samples, their volatile compounds were detected by GC-MS, and the content of identified volatiles were calculated and compared. A total of 57

Pu-erh teas were analysed, and the TICs of three different types of Pu-erh tea were shown in Figure S1 (see electronic supplementary information). The identified volatile compounds and their relative contents (%) (mean value and standard deviation) were summarized in Table 2. A total of 122 aroma compounds were identified in all tea samples: 116 components were identified in 10 aged teas, 82 components were identified in 23 raw teas, and 105 components were identified in 24 ripe teas.

The volatile components identified in 57 Pu-erh teas mainly included hydrocarbons, alcohols, ketones, methoxyphenolic compounds, esters, aldehydes, furans, nitrogen compounds, lactones, and acids. As listed in Table 2, obvious differences in the volatile composition and content among these three different Pu-erh teas can be found. The main volatile components in 10 aged teas were hexadecanoic acid, dihydroactinidiolide, caffeine, linalool, 6,10,14-trimethyl-2-pentadecanone, β-ionone, cedrol, and phytol; the main volatile components in 23 raw teas were linalool, tridecane, caffeine, dihydroactinidiolide, β-ionone, 6,10,14-trimethyl-2-pentadecanone, dodecane, etc.; and the main volatile components in 24 ripe teas were 1,2,3-trimethoxybenzene, hexadecanoic acid, 1,2,4-trimethoxybenzene, dihydroactinidiolide, 6,10,14-trimethyl-2-pentadecanon, caffeine, and 1,2,3-trimethoxy-5-methyl-benzene. The Pu-erh raw tea retains the flavour and colour of the original tea because of the lack of post fermentation in the manufacturing process, and therefore, it is not surprising that its aroma components and taste were similar to other green teas, such as Longjing tea [25], Hubei green tea [16], Biluochun green tea [14], etc. The most obvious feature was that the content of alcohols accounted for a higher proportion (30.65%) in raw teas, especially linalool, phytol, and geraniol. These volatiles, together with some other terpene alcohols, were likely the major contributor to the herbaceous and pleasant flowery odour of the raw tea [28]. In contrast, the content of these volatile compounds in the aged tea and ripe tea were only 21.95% and 14.69%, respectively. Because the aged teas were not processed by fast pile fermentation, the relative content of alcohols is obviously lower than that of the raw tea. There seems to be two possible reasons for such a decrease: the first was that some lower boiling alcohol compounds were volatilized during long-term storage process and the second was that some alcohol compounds were transformed into other substances by exogenous microbes and natural oxidation during long-term storage. In the case of the ripe tea, the fermentation process may effectively facilitate the oxidation and decomposition of some alcohols and therefore lead to the lowest content of alcohols among the three types of teas. The content of hydrocarbons in raw tea was higher than in the aged and ripe teas, especially that of tridecane and dodecane. This is possibly a result of the effective preservation of these components (that are enriched in the fresh tea leaves) in green tea without specific processing. Most of these identified hydrocarbons were saturated hydrocarbons; they usually make a minor contribution to tea flavour [29]. Unsaturated hydrocarbons play a vital role in the aroma of the tea; their content was relatively low, e.g., only 2.56% in ripe tea, and as such, may make a limited contribution to the aroma of Pu-erh tea. Ketones, with a sweet and woody flavour, exhibit comparable content in the three types of Pu-erh. We speculate that the processing and storage time has only a minor effect on the content of ketone compounds, and the contribution of ketone compounds to aromas of different types of Pu-erh tea is insignificant, similar to ionone. In addition, the content of methoxyphenolic compounds showed a significant difference ($p < 0.001$) among these three types of Pu-erh tea. The content of methoxyphenolic compounds was found to be the highest (27.78%) in ripe teas in comparison with the aged teas (11.85%) and raw teas (1.05%). Methoxyphenolic compounds made a strong contribution to the stale flavour of Pu-erh ripe teas, and as such, can be used as potential markers to distinguish among these three types of Pu-erh tea [26]. Some methoxyphenolic components, such as 1,2,3-trimethoxybenzene and 1,2,4-trimethoxybenzene, were commonly found in the aged and ripe teas, demonstrating that these two teas share certain chemical similarities to each other. In fact, although the aged tea has not experienced a post fermentation process, its quality characteristics were similar to the ripe tea after lengthy ageing. Therefore, we suggested that the methoxyphenolic compounds could be the important factor causing similar quality characteristics between aged tea and ripe tea. Similarly, under the action of microorganisms and hot and humid conditions, Pu-erh ripe tea generates many methoxyphenolic compounds within a short time. Although microorganisms have an important role in the formation of methoxyphenolic compounds, the observed high content of methoxyphenolic compounds in the aged tea suggested that the formation of these compounds did not necessarily rely on microorganisms. Consequently, the formation mechanism of methoxyphenolic compounds needs to be investigated further. The content of aldehyde compounds in aged tea (3.12%) was relatively higher than in raw tea (1.55%) and ripe tea (1.08%). Among them, the content of (E)-2-decenal was the highest in the aged tea, whereas it was almost undetectable in the other two tea types. 2-Dodecenal was only detected in the aged tea, and the contents of hexanal and benzaldehyde were higher in the aged tea than in the other two tea types. Therefore, these volatiles are likely to make a notable contribution to the aroma of aged tea. With regard to the lactone compounds, only dihydroactinidiolide has been detected in all teas. Dihydroactinidiolide, which has a coumarin and musk flavour and has a certain contribution to the aroma of Pu-erh tea, may be a degradation product of β-carotene. As for the acid compounds, only hexadecanoic acid has been detected. Of note, its content was higher in aged tea (8.44%) and ripe tea (10.52%) than in raw tea (1.82%). It therefore seems likely that the content of hexadecanoic acid is related to the fermentation processing and storage time. In addition, the similar quality characteristics of aged tea and ripe tea may also be attributed to the comparable contents of hexadecanoic acid in these two tea types. The main nitrogen compound in the three types of Pu-erh tea was caffeine, which was mainly related to the taste characteristics of the tea. The contents of esters, phenolics and oxygen-containing heterocyclic compounds were low in all Pu-erh teas. 2-Pentyl furan, whose formation is related to the Maillard and Strecker degradation reactions of amino acids and sugars, was detected but showed significant differences in content among the three types of Pu-erh tea [30].

Pu-erh raw tea has an even richer set of chemical substances than regular green tea, including water-extractable substances and tea polyphenols, which provide a favourable material base for the transformation of chemical constituents during the post-fermentation process and the natural ageing process [2]. In terms of the volatile components in Pu-erh teas, after post-fermentation, alcohol and hydrocarbon component content reduced sharply, while methoxyphenolic components increased significantly, and as a result, a great change in aroma quality occurred. As mentioned above, it has been reported that longer ageing improves the quality of Pu-erh tea. It can be seen from our results that aged tea and ripe tea have some similarities based on aroma components, such as methoxyphenolic and alcohol compounds; this makes these two tea types share some similar aroma characteristics. Some water-soluble ingredients, such as tea polyphenols and flavonoid compounds, should be further compared to explore the similarities and differences of their taste characteristics. In a word, the aged tea was piled in a natural way without being processed by pile fermentation, and therefore, it is not surprising that some differences are observed in the content and composition of aroma components between the aged tea and the ripe tea. With the aid of microorganisms, ripe tea achieves similar quality characteristics as aged tea through rapid fermentation in a short time; these quality characteristics have been widely recognized by consumers.

In the process of long-term ageing, Pu-erh aged tea experiences complex chemical changes, resulting in a sharp decrease in levels of low-boiling alcohols and hydrocarbons, and an increase in some of the high-boiling acids, e.g., hexadecanoic acid. Because of storing the tea for several decades, some low-boiling point substances were lost via evaporation, while some ingredients were enriched because of chemical transformation from other compounds such as tea polyphenols. Du [24] and Lv [26] found that the formation of methoxyphenolic compounds may have a particularly close connection with the methylation of tea catechins. Overall, the mechanism of post fermentation and long-term ageing of Pu-erh tea needs to be studied in-depth to explore the effects of various conditions on changes of tea inclusions, including polyphenols and volatile components.

However, it is extremely difficult to predict the volatile change during the storage process of Pu-erh tea. Because tea aroma component is a very complicated system and easily influenced by many factors, it needs to simultaneously characterize large numbers of volatiles in data matrices. In subsequent work, study of the dynamic change of the chemical composition and content associated with processing and storage length of Pu-erh tea is necessary. Additionally, not all volatile components contribute to fragrance equivalently; sometimes the aromatic contribution of specific volatile ingredients on a large scale was required to study more techniques in order to investigate and expose their hidden characteristics. Therefore, electronic nose (e-nose) and gas chromatography-olfactometry (GC-O) can be used in further studies for sensory evaluation of Pu-erh teas.

Optimization of data scaling methods in PCA

PCA, a well-known unsupervised technique, has been proven to be a powerful tool in summarizing and further explaining large data sets both statistically and visually [31-34]. Of importance, the multidimensional data set can be transformed into 2D or 3D coordinates via principal component projection. At present, five automatic scaling procedures are available in SIMCA-P12 software: UV scaling (variables are centred and divided by their standard deviation), UVN scaling (same as UV, but the variable is not centred), Par scaling (variables are centred and divided by the square root of their standard deviation), ParN scaling (same as Par, but variables are not centred), and Ctr scaling (variable is centred but not scaled). UV scaling gives each variable an equal chance of being expressed in the statistical analysis, while Ctr scaling influences a variable that is related to its amplitude; hence, low content elements have little influence. Par scaling is a compromise between them; the influence of low amplitude variables is enhanced, and centring the variables before scaling is helpful to reduce the distortion of results induced by multicollinearity [35]. Thus, it was essential to determine which scaling method gives the most reliable and unbiased results for our research. For this reason, it was decided to compare (by means of PCA) the influence of all five scaling methods based on the content of identified volatiles in all Pu-erh teas, and the corresponding results were shown in Figure 1.

As shown in Figure 1a and 1c, raw teas and ripe teas are well separated, and old teas are located between ripe teas and raw teas. However, in UVN (Figure 1b), ParN (Figure 1d) and Ctr (Figure 1e), the separation between samples from the same or different processing methods was not clear, and samples from different groups overlapped. Moreover, samples from one group were highly scattered and were not clearly clustered. In addition, compared with the UV and Par models, data from old teas were scattered in UV; the value of t [1] and t [2] was larger than the Par model. Therefore, Par scaling is more suitable for the comprehensive analysis of Pu-erh tea volatile components.

Although PCA is an unsupervised method, it yielded sufficient results for most of the analysed types of Pu-erh teas (Figure 1c). The best discrimination was observed between Pu-erh raw teas and ripe teas. The overlap observed between old teas and ripe teas suggests that they have some similar chemical characteristics.

Difference in Aroma Compositions between Pu-erh Teas with Different ages

Because the collected teas have different ages in this study, it is interesting to probe the differences in volatile components between them with different ages. This analysis was performed through a PCA loading plot. The loading plots of different ages of old tea (a), raw tea (b) and ripe tea (c) were shown in Figure S2 (see electronic supplementary information). Figure S2a shows that there are differences in some aroma components among aged teas with different ages. These aromatic compounds mainly included hexadecanoic acid (V118), V98 (dihydroactinidiolide), V112 (caffeine), linalool (V28), 6,10,14-trimethyl-2-pentadecanone (V113), *β*-ionone (V90), cedrol (V102), phytol (V122), α-terpineol (V40), geranyl acetone (V86), 1,2,3-trimethoxybenzene (V62), etc. The content of these components was relatively higher, and coefficients of variance were larger among these aged teas; thus, they were relatively dispersed in the loading plot. In Figure S2b, some of the aroma components might cause differences among raw tea from different years. These aromatic compounds mainly included linalool (V28), tridecane (V60), V98 (dihydroactinidiolide), V112 (caffeine), *β*-ionone (V90), etc. Similar to the above analysis, the content of these components was higher in the raw tea. In Figure S2c, 1,2,3-trimethoxybenzene (V62), hexadecanoic acid (V118), V98 (dihydroactinidiolide), 1,2,4-trimethoxybenzene (V68), V112 (caffeine), and 6,10,14-trimethyl-2-pentadecanone (V113) might cause differences among ripe tea from different years. The content of these components was higher in the ripe tea.

CA analysis

Like PCA, cluster analysis (CA) is also an unsupervised data analysis method that requires no prior knowledge of the test sample [36]. CA is another method that we have adopted to extract information on differences among different Pu-erh teas. It divides all samples into groups (clusters) according to similarities and finds the similarity among objects in a multidimensional space, forming clusters between the nearest objects [37]. To establish the clusters, Ward's method was used as the amalgamation rule and the squared Euclidean distance was used as the metric [38]. The dendrogram results of the CA were shown in Figure 2, showing that these Pu-erh tea samples were clustered into three groups, which is similar to the results of PCA. However, we found that two old teas were gathered in the ripe teas cluster; they had some similar aroma characteristics, but sometimes when the similarities were very prominent, they were not enough for a highly sustained conclusion.

The results of PCA and CA may lead to an inaccurate conclusion because of the lack of sufficient information in the original data set and the inherent flaws of these two methods, i.e., data dimension reduction and unsupervised recognition. Therefore, it is difficult to discriminate the most important variables from this loading plot in order to obtain a more accurate classification model. An alternative method, supervised OPLS-DA, was adopted to reveal the most important variables among 122 variables but also to have better discrimination among different kinds of Pu-erh teas.

OPLS-DA of three different types Pu-erh teas

In PCA analysis, original variables are preserved as much as possible in the first few components, which may lead to poor separation of the groups when the variability between groups is less than that within groups. Alternatively, OPLS-DA is a supervised method that reveals the direct correlation between variables and categories with a linear regression model [39, 40]. It is often used to sharpen the partition between groups of observations and maximise the separation among classes [41]. The OPLS-DA data set is the same as in the PCA analysis. The model showed one orthogonal component, with R2Y=0.92% and Q2=0.77% in Pu-erh tea samples (Figure 3). It also reveals three significant classifications with different colours, although the classifications were observed to be scattered among the aged teas. The changes of aroma components can be affected by environmental conditions, such as temperature, humidity, and microorganisms. As a result, during long-term storage, tea aroma components could undergo complicated changes, and the difference in storage duration could lead to different aroma components of these aged teas.

Although the number of tea samples is limited, these results showed that it is possible to discriminate and classify different processing types of Pu-erh tea based on the analysis of the volatile contents using pattern recognition techniques such as PCA, CA, and OPLS-DA. In the present study, the number of aged tea samples is relatively few because of the difficulty in obtaining reliable old tea sources. Future studies will collect more standard samples for aged tea with different ages and focus on the effect of storage time on dynamic changes in ingredients because of biochemistry and the impact of different environmental conditions on different chemical components, ultimately providing a theoretical basis for the scientific storage of Pu-erh tea. In conclusion, our study can lay a foundation for improving the scientific value of Pu-erh tea and provide understanding of the chemical composition and differences of different processing types of Pu-erh tea for consumers.

Conclusions

In the present study, the aroma characteristics from different manufacturing types and ageing lengths of Pu-erh teas were investigated by using GC-MS combined with a chemometrics method. A total of 122 volatile components were identified, among which 116 compounds were from aged teas, 82 were from raw teas, and 105 were identified in ripe teas. Large differences in aroma components among three types of Pu-erh tea were observed. The characteristic volatiles in aged teas were hexadecanoic acid, dihydroactinidiolide, caffeine, linalool, 6,10,14-trimethyl-2-pentadecanone, *β*-ionone, cedrol, and phytol; the characteristic volatiles in raw teas were linalool, tridecane, caffeine, dihydroactinidiolide, *β*-ionone, 6,10,14-trimethyl-2-pentadecanone, dodecane, etc.; and the characteristic volatiles in 24 ripe teas were 1,2,3-trimethoxybenzene, hexadecanoic acid, 1,2,4-trimethoxybenzene, dihydroactinidiolide, 6,10,14-trimethyl-2-pentadecanon, caffeine, and 1,2,3-trimethoxy-5-methyl- benzene. The observed large changes in methoxyphenolic compounds confirm that they play the dominant role in unfermented Pu-erh tea during long-term storage, suggesting that these compounds could be used as an index for discriminating between aged tea and raw tea. PCA, CA and OPLS-DA performed well in distinguishing among different Pu-erh tea samples. In summary, we have demonstrated that a multivariate statistical method is a useful tool for analysing the compositional pattern of Pu-erh teas that were produced using different manufacturing types or that were obtained from different ageing lengths.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No.31460228, 31340073) and Science Research Foundation of Yunnan Province Education Department (No.2014Y089) and the Six Talents Peak Projects of Jiangsu Province (2014-WSN-007).

References

1. L. Du, C. Wang, J. Li, D. Xiao, C. Li and Y. Xu, *J. Agric. Food Chem.*, 2012, **61**, 561-568.
2. H.-p. Lv, Y.-j. Zhang, Z. Lin and Y.-r. Liang, *Food Res. Int.*, 2013, **53**, 608-618.
3. L. Zhang, N. Li, Z.-Z. Ma and P.-F. Tu, *J. Agric. Food Chem.*, 2011, **59**, 8754-8760.
4. S. Ahmed, U. Unachukwu, J. R. Stepp, C. M. Peters, C. Long and E. Kennelly, *J. Ethnopharmacol*, 2010, **132**, 176-185.
5. G. Xie, M. Ye, Y. Wang, Y. Ni, M. Su, H. Huang, M. Qiu, A. Zhao, X. Zheng, T. Chen and W. Jia, *J. Agric. Food Chem.*, 2009, **57**, 3046-3054.
6. Q. Chen, J. Zhao and S. Vittayapadung, *Food Res. Int.*, 2008, **41**, 500-504.

7. S. Cui, J. Wang, L. Yang, J. Wu and X. Wang, *J. Pharmaceut. Biomed. Anal.*, 2015, **102**, 64-77.
8. S.-B. Wu, R. S. Meyer, B. D. Whitaker, A. Litt and E. J. Kennelly, *J. Chromatogr. A.*, 2013, **1314**, 154-172.
9. M. Pérez-González, J. J. Gallardo-Chacón, D. Valencia-Flores and V. Ferragut, *Food Anal. Methods*, 2015, **8**, 612-623.
10. E. Xu, J. Long, Z. Wu, H. Li, F. Wang, X. Xu, Z. Jin and A. Jiao, *J. Food Sci.*, 2015, **80**, C1476-C1489.
11. T. Ikeda, S. Kanaya, T. Yonetani, A. Kobayashi and E. Fukusaki, *J. Agric. Food Chem.*, 2007, **55**, 9908-9912.
12. W. Pongsuwan, E. Fukusaki, T. Bamba, T. Yonetani, T. Yamahara and A. Kobayashi, *J. Agric. Food Chem.*, 2007, **55**, 231-236.
13. S.-B. Wu, K. Dastmalchi, C. Long and E. J. Kennelly, *J. Agric. Food Chem.*, 2012, **60**, 7513-7525.
14. S.-D. Lv, Y.-S. Wu, Y.-Z. Song, J.-S. Zhou, M. Lian, C. Wang, L. Liu and Q.-X. Meng, *Food Anal. Methods*, 2015, **8**, 321-333.
15. J. Lin, P. Zhang, Z. Pan, H. Xu, Y. Luo and X. Wang, *Food Chem.*, 2013, **141**, 259-265.
16. N. Ye, L. Zhang and X. Gu, *Food Anal. Methods*, 2012, **5**, 856-860.
17. J.-E. Lee, B.-J. Lee, J.-O. Chung, H.-J. Shin, S.-J. Lee, C.-H. Lee and Y.-S. Hong, *Food Res. Int.*, 2011, **44**, 597-604.
18. M. C. A. Marcelo, C. A. Martins, D. Pozebon, V. L. Dressler and M. F. Ferrão, *Microchem. J.*, 2014, **117**, 164-171.
19. Y. He, X. Li and X. Deng, *J. Food Eng.*, 2007, **79**, 1238-1242.
20. K. Fraser, G. A. Lane, D. E. Otter, S. J. Harrison, S.-Y. Quek, Y. Hemar and S. Rasmussen, *Food Chem.*, 2014, **151**, 394-403.
21. K. M. Ku, J. Kim, H.-J. Park, K.-H. Liu and C. H. Lee, *J. Agric. Food Chem.*, 2009, **58**, 345-352.
22. L. Xu, D.-H. Deng and C.-B. Cai, *J. Agric. Food Chem.*, 2011, **59**, 10461-10469.
23. S. Lv, Y. Wu, J. Zhou, M. Lian, C. Li, Y. Xu, S. Liu, C. Wang and Q. Meng, *Plos one*, 2014, **9**, e116428.
24. L. Du, J. Li, W. Li, Y. Li, T. Li and D. Xiao, *Food Res. Int.*, 2014, **57**, 61-70.
25. J. Lin, Y. Dai, Y.-n. Guo, H.-r. Xu and X.-c. Wang, *J. ZheJiang Univ. Sci. B.*, 2012, **13**, 972-980.
26. H.-P. Lv, Q.-S. Zhong, Z. Lin, L. Wang, J.-F. Tan and L. Guo, *Food Chem.*, 2012, **130**, 1074-1081.
27. S.-D. Lv, Y.-S. Wu, J.-S. Zhou, M. Lian and Q.-X. Meng, *J. Chem. Pharm. Res.*, 2014, **6**, 246-253.
28. L.-F. Wang, J.-Y. Lee, J.-O. Chung, J.-H. Baik, S. So and S.-K. Park, *Food Chem.*, 2008, **109**, 196-206.
29. C. Alasalvar, B. Topal, A. Serpen, B. Bahar, E. Pelvan and V. Gökmen, *J. Agric. Food Chem.*, 2012, **60**, 6323-6332.
30. Z. Qin, X. Pang, D. Chen, H. Cheng, X. Hu and J. Wu, *Food Res. Int.*, 2013, **53**, 864-874.
31. S. Tianniam, L. Tarachiwin, T. Bamba, A. Kobayashi and E. Fukusaki, *J. Biosci. Bioeng.*, 2008, **105**, 655-659.

32. H. Yamamoto, H. Yamaji, Y. Abe, K. Harada, D. Waluyo, E. Fukusaki, A. Kondo, H. Ohno and H. Fukuda, *Chemometr. Intell. Lab. Syst.*, 2009, **98**, 136-142.
33. J. Zheng, R. Liang, C. Wu, R. Zhou and X. Liao, *Food Res. Int.*, 2014, **56**, 77-84.
34. N. Liu, Y. Liang, J. Bin, Z. Zhang, J. Huang, R. Shu and K. Yang, *Food Anal. Methods*, 2014, **7**, 472-480.
35. Ł. Zieliński, S. Deja, I. Jasicka-Misiak and P. Kafarski, *J. Agric. Food Chem.*, 2014, **62**, 2973-2981.
36. G. L. Monferrere, S. M. Azcarate, M. Á. Cantarelli, I. G. Funes and J. M. Camiña, *J. Food Sci.*, 2012, **77**, C1018-C1022.
37. G. E. Morlock, P. Ristivojevic and E. S. Chernetsova, *J. Chromatogr. A.*, 2014, **1328**, 104-112.
38. C. Sârbu, R. D. Naşcu-Briciu, A. Kot-Wasik, S. Gorinstein, A. Wasik and J. Namieśnik, *Food Chem.*, 2012, **130**, 994-1002.
39. Y.-K. Kwon, Y.-S. Bong, K.-S. Lee and G.-S. Hwang, *Food Chem.*, 2014, **161**, 168-175.
40. X. Deng, Q. Liao, X. Xu, M. Yao, Y. Zhou, M. Lin, P. Zhang and Z. Xie, *Food Anal. Methods*, 2014, **7**, 1840-1847.
41. M. Wang, B. Avula, Y.-H. Wang, J. Zhao, C. Avonto, J. F. Parcher, V. Raman, J. A. Zweigenbaum, P. L. Wylie and I. A. Khan, *Food Chem.*, 2014, **152**, 391-398.

Figure captions

Figure 1 Influence of different scaling methods on PCA results: (a) UV; (b) UVN; (c) Par; (d) ParN; (e) Ctr, among “O” represented aged teas, “R” represented raw teas, “p” represented ripe teas.

Figure 2 CA results of all 57 Pu-erh teas

Figure 3 Score plot of 57 Pu-erh teas based on the content of volatiles

Table 1 The detail information of various Pu-erh teas used in this work (raw tea, ripe tea and aged tea)

No.	Sample ID	Type	Production year	Production period	Amount
1	O1	Aged tea	1914	100 years	1
2	O2	Aged tea	1984	30 years	1
3	O3	Aged tea	1989	25 years	1
4	O4	Aged tea	1991	23 years	1
5	O5	Aged tea	1994	20 years	1
6	O6	Aged tea	1997	17 years	1
7	O7	Aged tea	1998	16 years	1
8	O8	Aged tea	1999	15 years	2
9	O9	Aged tea	2000	11 years	1
10	R1	Raw tea	2004	10 years	1
11	R2	Raw tea	2005	9 years	1
12	R3	Raw tea	2006	8 years	1
13	R4	Raw tea	2007	7 years	1
14	R5	Raw tea	2008	6 years	1
15	R6	Raw tea	2009	5 years	3
16	R7	Raw tea	2010	4 years	3
17	R8	Raw tea	2011	3 years	2
18	R9	Raw tea	2012	2 years	2
19	R10	Raw tea	2013	1 year	8
20	P1	Ripe tea	1996	18 years	1
21	P2	Ripe tea	2000	14 years	1
22	P3	Ripe tea	2001	13 years	1
23	P4	Ripe tea	2002	12 years	1
24	P5	Ripe tea	2003	11 years	1
25	P6	Ripe tea	2004	10 years	2
26	P7	Ripe tea	2005	9 years	2
27	P8	Ripe tea	2006	8 years	2
28	P9	Ripe tea	2007	7 years	1
29	P10	Ripe tea	2008	6 years	1
30	P11	Ripe tea	2009	5 years	1
31	P12	Ripe tea	2010	4 years	2
32	P13	Ripe tea	2011	3 years	2
33	P14	Ripe tea	2012	2 years	3
34	P15	Ripe tea	2013	1 year	3

Table 2. Volatile compounds and their relative contents in three different types of Pu-erh teas

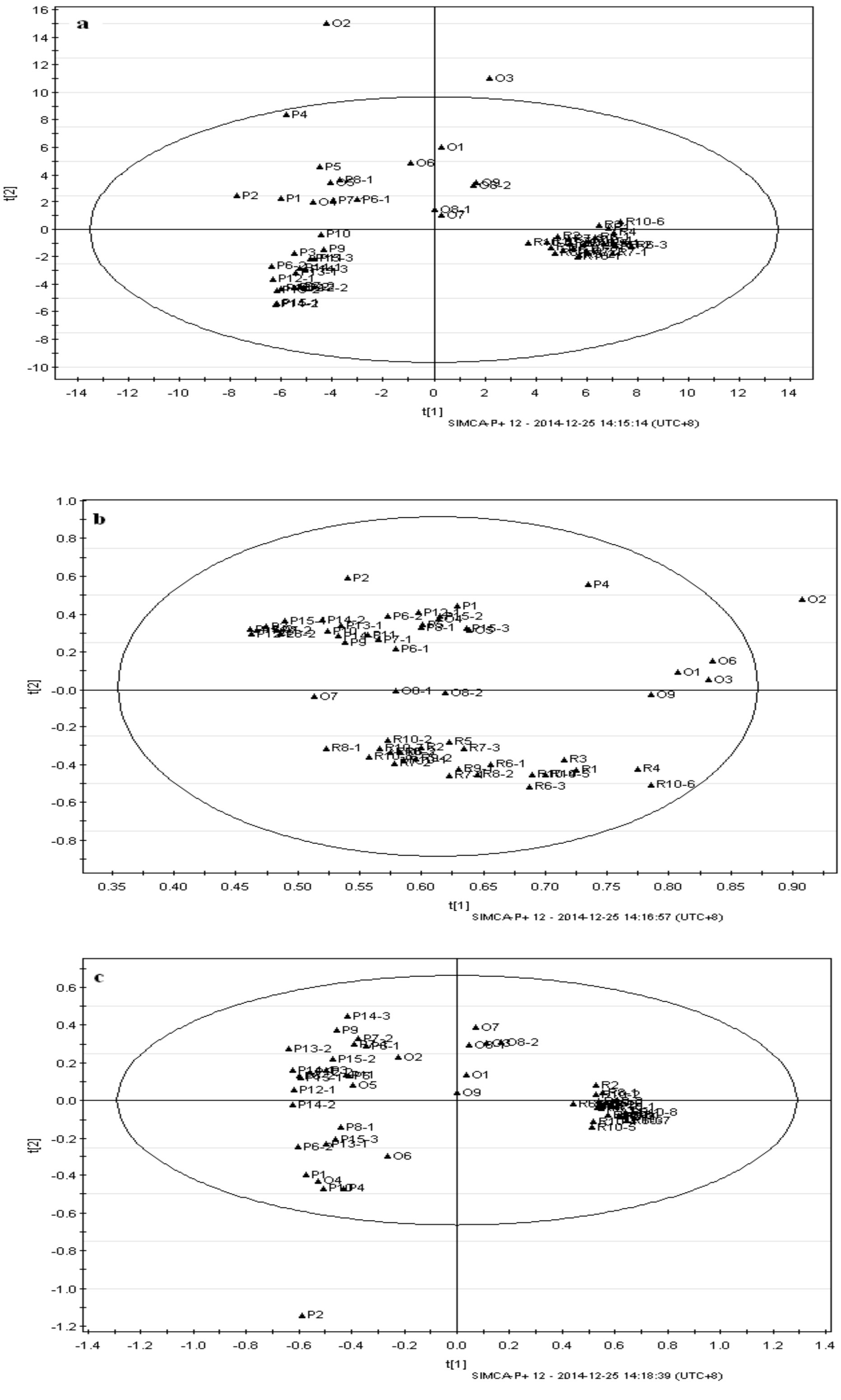
No.	RI[a]	Compound[b]	Relative percentage content[% (range)][c]		
			Aged tea (n=10)	Raw tea (n=23)	Ripe tea (n=24)
1	802	1-pentanol	0.00±0.00a	0.03±0.09a	0.00±0.00a
2	806	hexanal	0.26±0.29a	0.00±0.00b	0.07±0.12b
3	843	cis-3-hexenol	0.03±0.07a	0.00±0.00b	0.00±0.00b
4	861	1-hexyl alcohol	0.02±0.04 a	0.00±0.00 b	0.00±0.00b
5	884	2-heptanone	0.10±0.09a	0.03±0.05b	0.02±0.04b
6	894	2-heptanol	0.11±0.17a	0.02±0.04b	0.01±0.05b
7	957	benzaldehyde	0.29±0.17a	0.20±0.06b	0.14±0.08b
8	979	1-octen-3-ol	0.36±0.26a	0.90±0.99b	0.03±0.05a
9	982	2,3-octadione	0.04±0.05a	0.06±0.08b	0.01±0.02a
10	985	6-methyl-5-hepten-2-one	0.45±0.25a	0.25±0.11b	0.05±0.07c
11	989	2-pentyl-furan	0.49±0.25a	0.91±0.30b	0.21±0.13c
12	998	*cis*-2-(2-pentenyl)furan	0.09±0.09a	0.14±0.09a	0.03±0.05b
13	1010	α-terpinene	0.04±0.06a	0.12±0.06b	0.01±0.02a
14	1022	1-methyl-2-(1-methylethyl)-benzene	0.27±0.15a	0.19±0.18a	0.05±0.13b
15	1026	D-limonene	0.58±0.58a	1.20±0.79b	0.08±0.09c
16	1030	2-ethylhexanol	0.02±0.04a	0.00±0.00a	0.07±0.18a
17	1034	benzyl alcohol	0.14±0.16a	0.40±0.10b	0.02±0.04c
18	1037	*(E)*-3,7-dimethyl-1,3,6-octatriene	0.16±0.15a	0.27±0.09b	0.00±0.00c
19	1042	phenyl acetaldehyde	0.22±0.12a	0.24±0.10a	0.06±0.07b
20	1048	1-ethyl-1H-pyrrole-2-carbaldehyde	0.02±0.05a	0.07±0.11ab	0.14±0.16b
21	1051	Ocimene	0.05±0.08a	0.45±0.15b	0.00±0.00a
22	1056	γ-terpinene	0.30±0.19a	0.33±0.11a	0.07±0.17b
23	1064	acetophenone	0.07±0.07a	0.03±0.12a	0.03±0.06a
24	1068	*(E)*-2-octen-1-ol	0.19±0.15a	0.39±0.46b	0.00±0.00a
25	1072	linalool oxide I	0.90±0.61a	0.97±0.50a	0.82±0.53a
26	1087	linalool oxide II	1.48±1.24a	1.92±0.63a	1.65±0.87a
27	1092	*(E,E)*-3,5-octadien-2-one	0.23±0.47a	0.06±0.09b	0.12±0.16ab
28	1098	linalool	4.34±2.71a	14.72±4.51b	0.77±0.65c
29	1101	hotrienol	1.20±0.66a	1.76±0.48b	0.31±0.32c
30	1110	phenylethyl alcohol	0.25±0.17a	0.00±0.00b	0.25±0.23a
31	1135	benzene acetonitrile	0.64±0.75a	0.00±0.00b	0.13±0.32b
32	1137	2,5-pyrrolidinedione, 1-ethyl-	0.12±0.20a	0.02±0.08b	0.03±0.06b
33	1139	3-nonen-2-one	0.05±0.09a	0.00±0.00b	0.00±0.00b
34	1149	1,2-dimethoxybenzene	0.47±0.60a	0.01±0.03b	1.09±0.49c
35	1153	1,4-dimethoxy-2-methylbenzene	0.05±0.13a	0.00±0.00b	0.00±0.00c
36	1159	*(E)*-2-nonenal	0.09±0.14a	0.00±0.00b	0.06±0.10ab
37	1169	linalool oxide III	0.19±0.21a	0.00±0.00b	0.38±0.26c
38	1175	linalool oxide IV	1.20±0.58a	0.84±0.41a	1.31±0.85a
39	1178	naphthalene	0.45±0.40a	0.44±0.16a	0.49±0.59a
40	1188	α-terpineol	2.48±2.09a	2.53±1.18a	1.33±0.75b
41	1190	methyl salicylate	0.61±0.41a	0.68±0.76a	0.40±0.28a

42	1196	safranal	0.40±0.23a	0.37±0.10a	0.18±0.13b
43	1200	dodecane	0.32±0.85a	3.31±0.93b	0.03±0.06a
44	1205	decanal	0.27±0.27a	0.00±0.00b	0.26±0.14a
45	1218	*β*-cyclocitral	0.42±0.15a	0.67±0.18b	0.16±0.10c
46	1221	2,3-dihydrobenzofuran	0.25±0.65a	0.00±0.00a	0.26±0.86a
47	1224	3-carene	0.09±0.11a	0.21±0.08b	0.00±0.00c
48	1228	nerol	0.19±0.20a	0.29±0.13b	0.05±0.09c
49	1236	2-methoxy-4-methyl-1-(1-methylethyl)-benzene	0.20±0.18a	0.00±0.00b	0.04±0.10b
50	1241	3,4-dimethoxytoluene	0.85±0.94a	0.00±0.00b	0.88±0.79a
51	1256	geraniol	1.20±0.48a	1.59±0.46b	0.44±0.28c
52	1259	7-methoxybenzofuran	0.06±0.15a	0.00±0.00b	0.01±0.04b
53	1261	2-methoxybenzyl alcohol	0.10±0.11a	0.00±0.00b	0.03±0.06b
54	1263	*(E)*-2-decenal	1.07±2.06a	0.00±0.00b	0.01±0.05b
55	1265	3,5-dimethoxytoluene	1.24±3.51a	0.00±0.00b	0.07±0.11b
56	1285	2-methyl-naphthalene	0.52±0.36a	0.34±0.08b	0.37±0.28ab
57	1287	tridecene	0.00±0.00a	0.21±0.12b	0.00±0.00a
58	1290	indole	0.06±0.11a	0.06±0.16a	0.00±0.00a
59	1294	2-undecanone	0.29±0.29a	0.00±0.00b	0.20±0.20a
60	1300	tridecane	0.01±0.03a	8.05±2.06b	0.01±0.20a
61	1302	1-methyl-naphthalene	0.25±0.19a	0.00±0.00b	0.22±0.11a
62	1316	1,2,3-trimethoxybenzene	2.94±4.10a	0.34±0.19a	11.52±5.75b
63	1325	4-ethyl-1,2-dimethoxy-benzene	1.43±2.01a	0.01±0.04b	2.27±1.56a
64	1334	2,6,6-trimethyl-1-cyclohexene-1-ethanol	0.16±0.19a	0.00±0.00b	0.14±0.23a
65	1351	2,6-dimethoxyphenol	0.53±0.70a	0.31±0.18a	0.29±0.17a
66	1362	2-dodecenal	0.08±0.17a	0.00±0.00b	0.00±0.00b
67	1366	*α*-ylangene	1.51±1.48a	0.00±0.00b	0.71±1.65ab
68	1375	1,2,4-trimethoxybenzene	0.83±1.06a	0.00±0.00a	4.87±3.37b
69	1381	*β*-damascenone	0.04±0.11a	0.02±0.09a	0.00±0.00a
70	1383	1-methoxy-4-propenyl-benzene	0.15±0.39a	0.00±0.00b	0.02±0.08b
71	1387	*α*-gurjunene	1.70±5.38a	0.00±0.00b	0.00±0.00b
72	1389	*β*-guaiene	0.00±0.00a	0.00±0.00a	0.38±0.36b
73	1397	*cis*-jasmone	0.00±0.00a	0.40±0.39b	0.24±0.23b
74	1400	tetradecane	0.63±0.53a	1.05±0.19b	0.31±0.20c
75	1404	1,2,3-trimethoxy-5-methyl-benzene	1.49±2.67a	0.01±0.05a	3.34±3.28b
76	1406	6,10-dimethyl-2-undecanone	0.12±0.25a	0.00±0.00b	0.00±0.00b
77	1408	1,2-dimethoxy-4-n-propenyl-benzene	0.10±0.33a	0.00±0.00a	0.33±0.70a
78	1411	*α*-cedrene	0.65±0.82a	0.15±0.25b	0.49±0.47a
79	1417	*β*-caryophyllene	0.21±0.44a	0.20±0.25a	0.02±0.10b
80	1428	α-ionone	1.12±0.48a	1.22±0.39a	0.88±0.37b
81	1433	1,2-benzopyrone	0.53±0.32a	0.43±0.19a	0.43±0.52a
82	1438	dihydro-*β*-ionone	0.13±0.21ab	0.00±0.00a	0.27±0.39b
83	1442	1-methoxy-naphthalene	0.07±0.12a	0.00±0.00a	0.52±0.25b
84	1447	2-methoxy-naphthalene	0.08±0.16a	0.00±0.00a	0.71±0.56b

85	1449	1,2,3,4-tetramethoxybenzene	0.71±0.80a	0.00±0.00b	1.12±0.62c
86	1455	geranyl acetone	2.25±1.22a	2.33±0.64a	1.51±0.62b
87	1460	aromandendrene	0.23±0.38a	0.00±0.00b	0.03±0.10b
88	1468	5-methoxy-6,7-dimethyl-benzofuran	0.55±0.65a	0.00±0.00b	0.09±0.20b
89	1483	4-(2,6,6-trimethylcyclohexa-1,3-dienyl)-but-3-en-2-one	1.44±2.07ab	0.25±0.17a	2.72±2.55b
90	1487	β-ionone	3.28±1.80a	5.05±1.23b	3.02±1.27a
91	1489	pentadecene	0.38±0.64a	0.89±0.27b	0.00±0.00c
92	1492	2-tridecanone	0.19±0.39a	0.00±0.00a	0.69±1.04b
93	1500	pentadecane	0.77±1.02a	0.62±0.24a	0.50±0.38a
94	1502	1,2-dimethoxy-4-(1-propenyl)benzene	0.00±0.00a	0.00±0.00a	0.34±0.35b
95	1504	cuparene	0.56±0.93a	0.00±0.00b	0.35±0.76ab
96	1506	dibenzofuran	0.48±0.57a	0.68±0.63a	0.56±0.30a
97	1508	α-farnesene	0.00±0.00a	1.40±0.86b	0.49±0.50c
98	1528	dihydroactinidiolide	5.74±3.50a	6.16±1.05a	4.37±1.38b
99	1549	1,2,3-trimethoxy-5-(2-allylbenzene)	0.15±0.34a	0.00±0.00b	0.00±0.00b
100	1554	nerolidol	0.05±0.14a	0.29±0.39a	1.37±1.60b
101	1572	fluorene	0.88±0.44a	0.97±0.32a	0.85±0.42a
102	1598	cedrol	3.57±2.45a	0.57±0.29b	1.73±1.57c
103	1600	hexadecane	1.31±0.80a	1.32±0.47a	1.48±0.89a
104	1653	α-cadinol	0.36±0.48a	0.96±0.23b	0.83±0.38b
105	1659	2,2',5,5'-tetramethyl-1,1'-biphenyl	0.22±0.22a	1.05±0.26b	0.46±0.36c
106	1664	2-methyl-hexadecane	0.10±0.22a	0.17±0.28ab	0.35±0.24b
107	1700	heptadecane	1.51±1.30a	0.87±0.70a	1.56±0.99a
108	1706	2,6,10,14-tetramethyl-pentadecane	1.77±1.14a	2.66±0.91b	1.78±1.16a
109	1765	anthracene	1.40±0.48a	0.64±0.42b	1.54±0.74a
110	1800	octadecane	0.94±0.78a	0.41±0.45b	1.06±0.73a
111	1809	2,6,10,14-tetramethyl-hexadecane	0.94±0.91ab	0.49±0.42a	1.07±0.90b
112	1840	caffeine	4.83±2.31ab	6.13±2.89b	4.05±1.90a
113	1846	6,10,14-trimethyl-2-pentadecanone	4.04±2.32a	3.75±1.90a	4.21±2.31a
114	1900	nonadecane	0.15±0.24a	0.13±0.20a	0.13±0.20a
115	1918	farnesyl acetone	0.29±0.26a	0.38±0.44a	0.41±0.31a
116	1927	hexadecanoic acid methyl ester	0.36±0.31a	0.51±1.70a	0.46±0.33a
117	1949	isophytol	0.13±0.15a	0.06±0.07a	1.03±0.64b
118	1975	hexadecanoic acid	8.44±3.16a	1.82±1.49b	10.52±5.43a
119	2000	eicosane	0.01±0.04a	0.05±0.08a	0.18±0.18b
120	2093	methyl linoleate	0.07±0.04ab	0.00±0.00a	0.08±0.14b
121	2099	methyl linolenate	0.17±0.13a	0.17±0.26a	0.31±0.34a
122	2122	phytol	3.28±4.20a	2.41±1.58a	2.12±0.34a
Alcohols			21.95a	30.65b	14.69a
Hydrocarbons			18.91a	28.19b	15.07a
Ketones			14.66a	14.26a	14.81a
Esters			1.21a	1.36a	1.25a
Aldehydes			3.12a	1.55b	1.08b

Nitrogen compounds	5.65ab	6.21b	4.21a
Lactones	5.74a	6.16a	4.37b
Methoxyphenolic	11.85a	1.05b	27.78c
Acids	8.44a	1.82b	10.52a
Others	1.36a	1.36a	0.79a

[a] RI, retention indices as determined on HP-5MS column using the homologous series of n-alkanes (C_8–C_{40}). [b] Compounds were listed in order of retention time. [c] The content of volatile compounds were represented as mean value ±standard deviation (mean ± SD), different letters indicated significant differences ($P < 0.05$, ANOVA, Duncan's multiple range test).


a
t[2]
t[1]
SIMCA-P+ 12 - 2014-12-25 14:15:14 (UTC+8)
O2
O3
P4
O1
O6
P5
P8-1
O9
O8-2
P2
P1
O4
P7
P6-1
O8-1
O7
R10-6
P10
P9
P3
P6-2
P12-1
P15-2
b
t[2]
t[1]
SIMCA-P+ 12 - 2014-12-25 14:16:57 (UTC+8)
P2
P4
O2
P1
P6-2
P15-2
P14-2
P13-1
P8-1
P15-3
P7-1
P6-1
O6
O1
O3
O7
O8-1
O8-2
O9
R8-1
R10-2
R5
R7-3
R3
R6-1
R1
R4
R8-2
R10-6
R6-3
c
t[2]
t[1]
SIMCA-P+ 12 - 2014-12-25 14:18:39 (UTC+8)
P14-3
O7
P9
P7-2
P8-1
O8-2
P13-2
P15-2
O2
O1
O5
P12-1
O9
R2
P14-2
R10-8
R10-5
P8-1
P15-3
P6-2
P13-1
O6
P1
O4
P4
P2

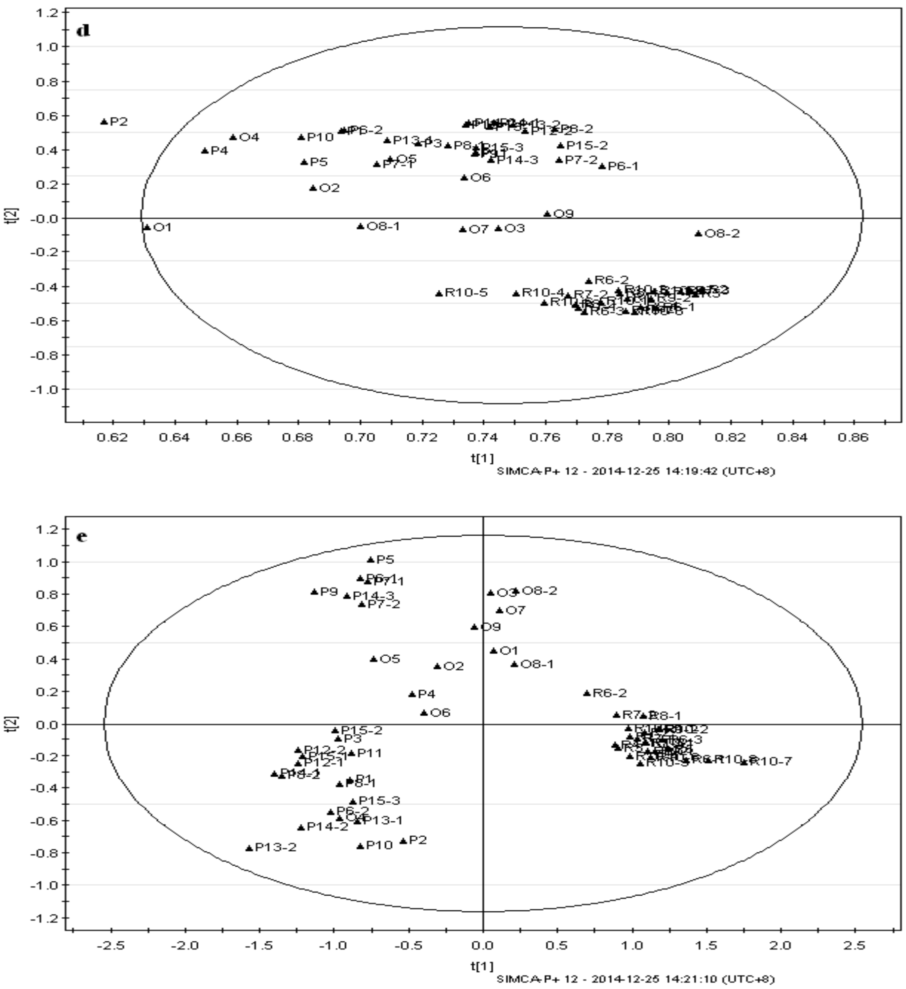


Figure 1 Influence of different scaling methods on PCA results: (a) UV; (b) UVN; (c) Par; (d) ParN; (e) Ctr, among “O” represented aged teas, “R” represented raw teas, “p” represented ripe teas.

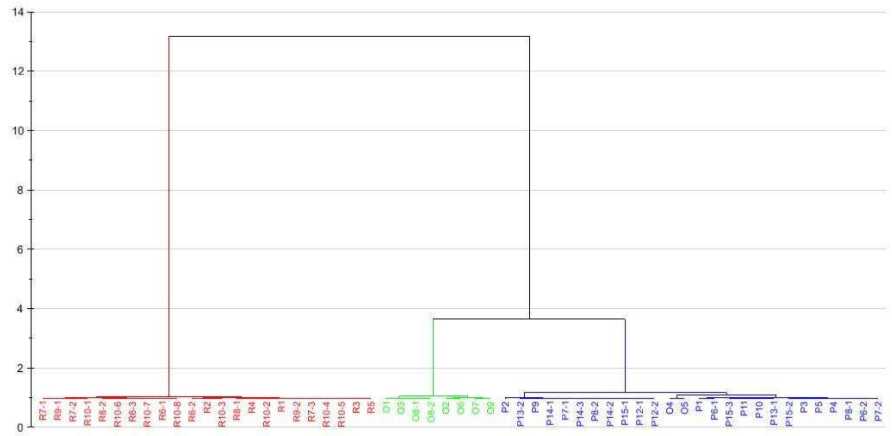


Figure 2 CA results of all 57 Pu-erh teas

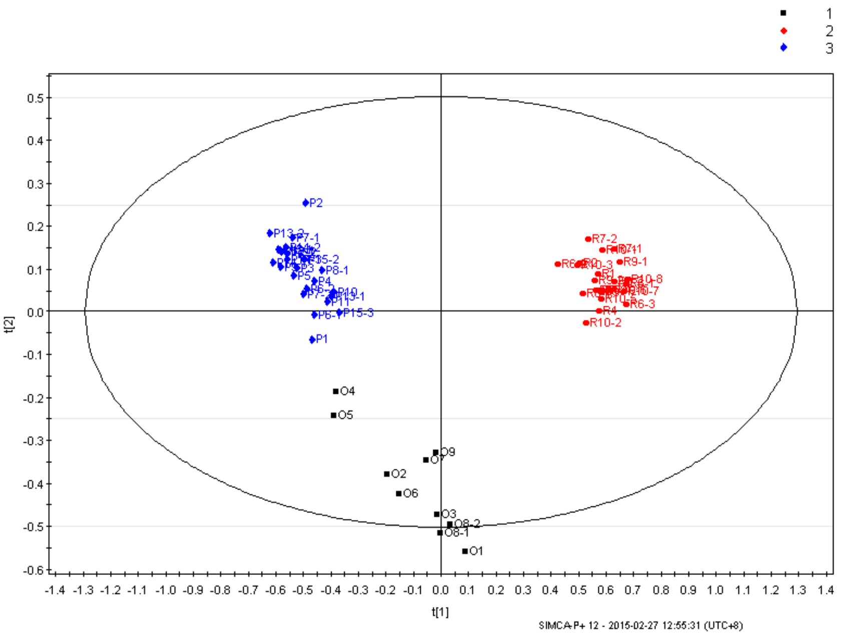
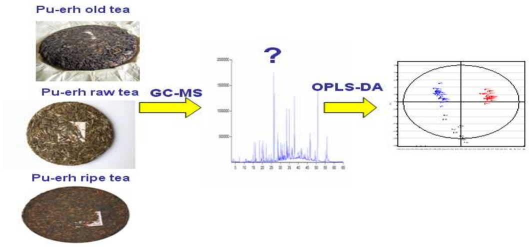


Figure 3 Score plot of 57 Pu-erh teas based on the content of volatiles



Graphical abstracts