Application Note

Food Testing & Agriculture

Study of the Metabolites and Flavor Characteristics in Different Subtypes of White Tea by Metabolomics Profiling

Authors

Chen Yang, Weidong Dai, Junfeng Tan, and Zhi Lin
Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, China

Meiling Lu
Agilent Technologies (China) Co., Ltd, Beijing, China

Abstract

This Application Note describes the influence of nonvolatile compounds in white tea on flavor characteristics through an UHPLC-Q-TOF/MS based nontarget metabolomics profiling approach. Profiling of the tea metabolome using UHPLC-Q-TOF/MS followed by feature extraction and alignment resulted in 1,915 metabolite features. Principal component analysis (PCA) and supervised partial least square differential analysis (PLSDA) based on above features demonstrate a clear separation of three subtypes of white tea samples. Up to 99 compounds were identified by matching against authentic standards and databases. Forty-one metabolites exhibit high correlation with flavor; theanine, aspartic acid, asparagine, and AMP are positively correlated with the umami flavor, and flavan-3-ols, theasinensins, procyanidin B3, and theobromine have positive correlations with higher bitterness and astringency flavors. The results demonstrate that metabolomic profiling can be an effective approach to differentiate tea characteristics through characteristic compounds, and that such compounds are potential markers for determining the artificial adulteration and mislabeling of white tea in the market.

Introduction

White tea (*Camellia sinensis* L.) is one type of traditional Chinese tea mainly produced in the north and east of the Fujian province in China. In recent years, white tea has been demonstrated to have potential beneficial health effects[1], promoting its consumption in China. White teas of different qualities differ in bio-activity, aroma, flavor, and commercial value. Few known highly abundant substances have been quantitatively investigated in white tea[2,3]. To further the understanding of white tea qualities, and prevent artificial adulteration and mislabeling, a metabolomics approach was applied to systematically study the nonvolatile components in the different subtypes of white tea, aiming to elucidate the characteristic and differential metabolites and their association with white tea characteristics in terms of umami, bitterness, and astringency.

Experimental

Sample preparation

Fresh tea leaves were classified into subtypes, including 10 Silver Needle (SN), eight White Peony (WP), and 11 Shoumei (SM). These tea leaves were processed into white tea in accordance with a typical white tea processing procedure, which includes withering and drying. Once white tea was produced, it was ground into powder with 100 mesh, and stored at 4 °C. One gram of tea powder from each sample was then infused with 100 mL of boiling water, and maintained at 100 °C for five minutes prior to filtration with a cellulose filter. The equivalent quantitation of tea flavor including umami, bitterness, and astringency flavor was described in detail in a previous study[4].

Tea metabolome extraction

Tea powder (0.1 g for each sample) was suspended in 10 mL of hot deionized water (100 °C) for five minutes to extract tea metabolites. A 2 mL amount of the solution was then centrifuged at 10,000 × g for 10 minutes. The supernatants were filtered through a 0.22 µm membrane, then analyzed by UHPLC-Q-TOF/MS. Quality control (QC) samples were prepared by mixing an equal volume of each tea sample (50 µL). The samples were used to evaluate the data reliability for metabolomics analysis. Figure 1 shows the schematic diagram for sample preparation.

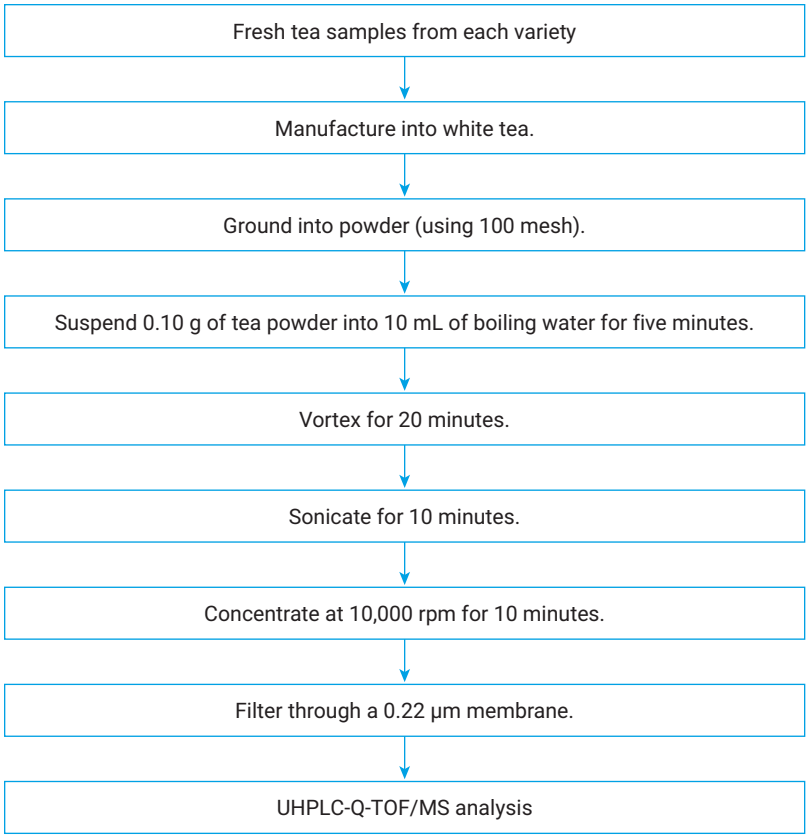

Figure 1. Procedure for tea sample preparation.

Nontarget metabolomics investigation

An UHPLC system coupled with a Q-TOF mass spectrometer was applied for data acquisition. Table 1 shows the detailed conditions for HPLC separation and MS detection. The TOF scanning and auto MS/MS data were acquired using 6540 LC/Q-TOF, and target MS/MS data for the standard compounds were acquired using 6545 LC/Q-TOF.

Workflow for metabolomics analysis

To conduct nontarget metabolomics profiling analysis, the accurate MS spectra for each group of samples and the QC samples were initially acquired in TOF scanning mode. The resulting raw data were subjected to molecular feature extraction using Agilent MassHunter Profinder software (Version 8.0), and the results were imported into Mass Profiler Professional (MPP) software (Version 14.8, Agilent Technologies, Santa Clara, CA) for peak alignment and preliminary filtration based on the variance of coefficients (≤30 %) for the peaks in the QC samples. The resulting peaks were identified by:

- Matching against databases, including Metlin and HMDB
- Matching with standard compounds, or
- Interpretation based on the MS/MS spectra

To demonstrate the contribution of the differential metabolites to the tea characteristics, the chemometric methods including principle components analysis, partial least square differential analysis, and hierarchy cluster analysis were applied. Pearson correlation analysis was also conducted between the abundance of the identified differential metabolites and the tea flavor. The resultant differential metabolites with high correlation were subjected to validation among all three groups of tea samples. Figure 2 presents the entire workflow for tea metabolomic profiling analysis.

Table 1. Instrument conditions.

Parameter	Value
LC Conditions	
HPLC	Agilent 1290 Infinity II LC with built-in degasser; autosampler with temperature control; column temperature control compartment
Column	Agilent ZORBAX Eclipse Plus C18, 150 × 3.0 mm, 1.8 µm
Column Temperature	40 °C
Mobile Phase	A) 0.1 % Formic acid in H_2O B) Methanol
Flow Rate	0.4 mL/min
Injection Volume	3.0 µL
Needle Backflush	5 seconds with pure methanol
Gradient Elution Profile	0 to 4 minutes 10 to 15 %B 4 to 7 minutes 15 to 25 %B 7 to 9 minutes 25 to 32 %B 9 to 16 minutes 32 to 40 %B 16 to 22 minutes 40 to 55 %B 22 to 28 minutes 55 to 95 %B 28 to 30 minutes 95 %B 30 to 31 minutes 95 to 10 %B 31 to 35 minutes 10 %B
ESI-Q-TOF MS Conditions	
MS	Agilent 6540/6545 ultrahigh-definition accurate-mass Q-TOF LC/MS with dual Jet Stream ESI
Polarity	Positive ionization
Drying Gas Temperature	300 °C
Drying Gas Flow Rate	8 L/min
Nebulizer Gas Pressure	35 psi
Sheath Gas Temperature	300 °C
Sheath Gas Flow Rate	11 L/min
Capillary Voltage	3,500 V
MS Scan Range	*m/z* 100 to 1,100
MS/MS Scan Range	*m/z* 50 to 1,100
Reference Ions	*m/z* 121.0509/922.0098
Scanning Mode	TOF scanning, autoMS/MS, and target MS/MS

Results and discussion

Separation and detection of tea extract

An optimized UHPLC gradient elution was applied to separate thousands of compounds in the white tea extract, as described in a previous report[5]. The eluate from the chromatographic column was directed to electrospray ionization-interfaced Q-TOF/MS for detection in scan mode. The selected condition enables reasonable separation of the unknown compounds, as shown in the typical total ion chromatograms for the QC sample (Figure 3). The acquired data were subjected to recursive molecular feature extraction (MFE) using MassHunter Profinder software (V 8.0) to obtain the reliable metabolite features.

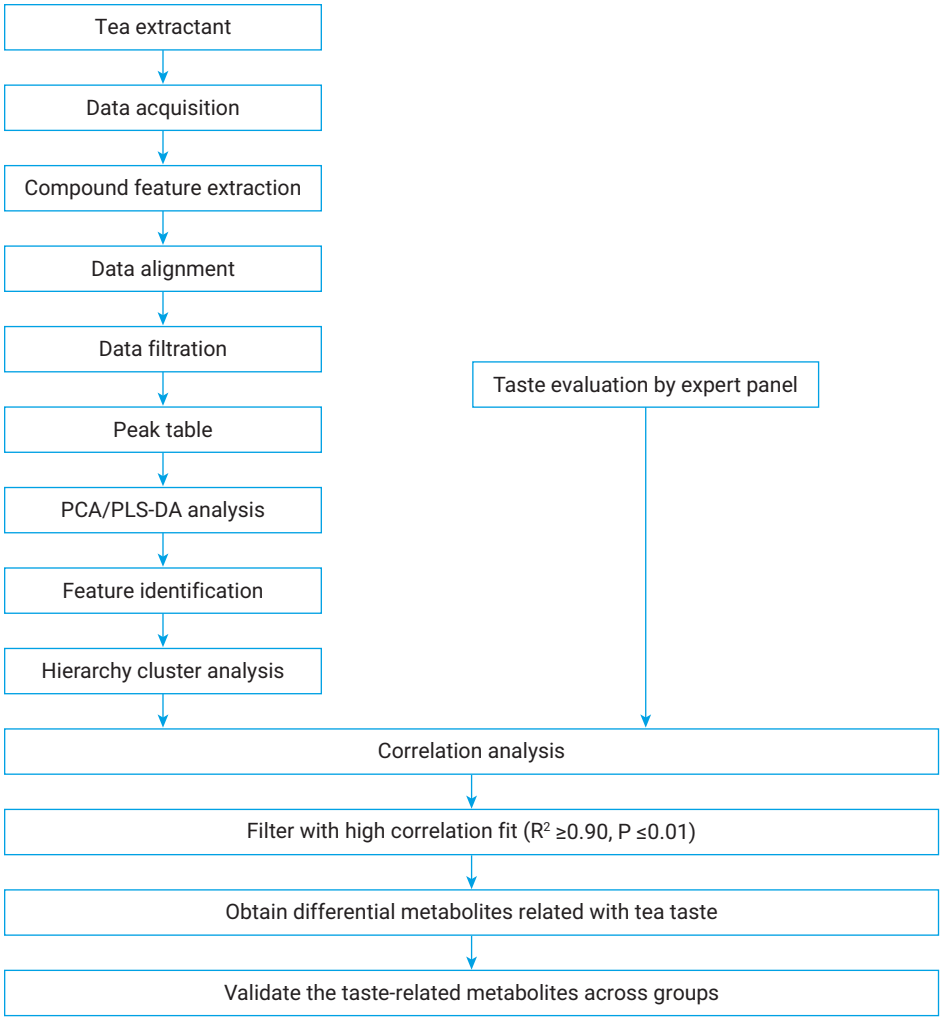

Figure 2. Schematic diagram showing the workflow for metabolic profiling analysis.

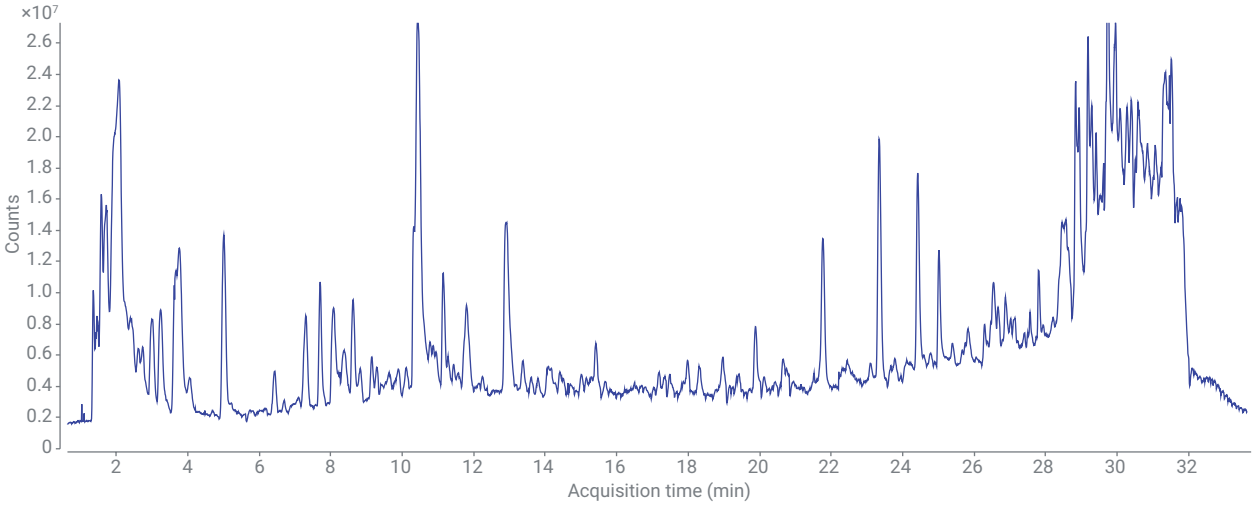

Figure 3. Typical total ion chromatograms for the QC sample of white teas.

Chemometric analysis

The resultant metabolite features from each sample were imported into MPP (V.14.8) software for features alignment and filtration with a variance of coefficients for all the features in the QC samples within 30 %. This resulted in a feature table with 1,915 compound features. Based on these features, the PCA score plot demonstrated that the difference in the metabolite patterns can separate the white tea samples by the category of: SN, WP, and SM (Figure 4A). A supervised PLSDA model was then established and validated for sample prediction based on the metabolites' patterns. The PLSDA score plot also exhibits a clear separation of samples by category based on the selected features (Figure 4B). The confusion matrix table demonstrates the excellent accuracy of the PLSDA model prediction (Table 2).

Table 2. The confusion matrix table showing the accuracy of the prediction.

	SN_p	WP_p	SM_p	Accuracy
(T)SN	10	0	0	100
(T)BP	0	8	0	100
(T)SM	0	0	11	100
Overall accuracy				100

T = true
p = prediction

Tentatively identified metabolite features

Metabolite features (99) were identified based on database searching and accurate MS/MS spectra interpretation. Some of them were further confirmed using the authentic standards shown in Table 3.

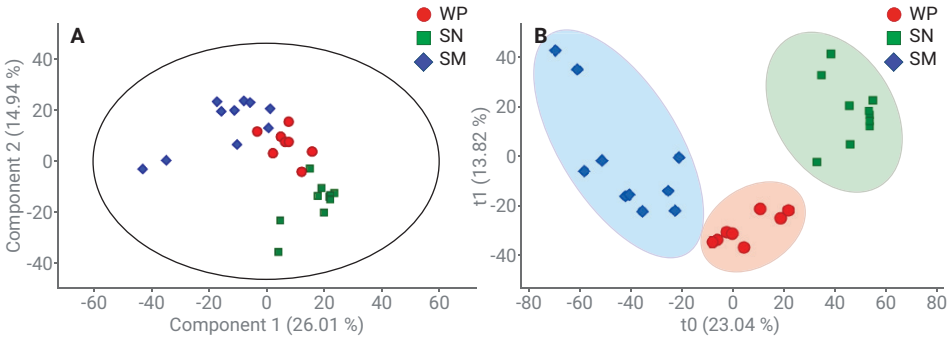

Figure 4. The PCA score plot (A) and the PLSDA score plot (B) demonstrating the separation of the three subtypes of white teas based on the extracted 1,915 metabolite features. Red circle: WP; green square: SN; blue diamond: SM.

Table 3. Identified compounds with retention time, accurate mass, mass error and their MS2 fragment.

No.	Identified Compound	RT (min)	Accurate Mass (*m/z*)	Adduct	Error (ppm)	MS2 Fragments
1	Caffeine*	10.53	195.0875	M+H	-1	138, 110, 69
2	Choline	1.49	104.107	M+H	-4.8	60, 58
3	Theobromine*	6.46	181.0718	M+H	-1.1	138, 110, 83
4	Betaine	1.58	118.0861	M+H	-1.7	72, 55
5	Glycerophosphocholine	1.75	258.1084	M+H	-6.6	184, 104, 60
6	Theanine*	2.07	175.1076	M+H	-0.6	158, 130, 84, 56
7	Pheylalanine(Phe)*	5.06	166.0855	M+H	-4.8	120, 103, 91, 77
8	Isoleucine(Ile)*	3.27	132.1017	M+H	-1.5	114, 86, 72, 56
9	Leucine (Leu)*	3.03	132.1017	M+H	-1.5	86, 69, 56
10	Proline (Pro)*	1.66	116.0708	M+H	1.7	70, 43
11	Tryptophan(Typ)*	7.79	205.097	M+H	-1	188, 159, 146, 118
12	Valine(Val)*	1.97	118.0862	M+H	-0.9	72, 63, 58, 55
13	Tyrosine(Tyr)*	2.74	182.0809	M+H	-1.7	136, 119, 91, 77
14	Asparagine (Asn)*	1.52	133.0606	M+H	-1.5	74
15	Glutamine (Gln)*	1.54	147.076	M+H	-2.7	130, 84, 56
16	Aspartic acid (Asp)*	1.55	134.045	M+H	1.5	88, 74, 43
17	*gamma*-Aminobutyric acid (GABA)*	1.52	104.0707	M+H	1	86, 69
18	Catechin (C)*	8.86	291.0865	M+H	0.7	139, 123, 95
19	Epicatechin (EC)*	11.32	291.0864	M+H	0.3	207, 139, 123, 55
20	Gallocatechin (GC)*	5.24	307.0815	M+H	0.7	223, 195, 163, 139
21	Epigallocatechin (EGC)*	8.65	307.0813	M+H	0	289, 153, 139
22	Epicatechingallate (ECG)*	13.27	443.0971	M+H	-0.5	273, 153, 139, 123
23	Gallocatechingallate (CG)*	11.63	459.0923	M+H	0.2	289, 181, 153, 139
24	Epigallocatechingallae (EGCG)*	10.43	459.0923	M+H	0.2	441, 289, 153, 139
25	Epigallocatechin digallate	13.45	611.103	M+H	-0.2	441, 289, 153
26	Epigallocatechin 3-methylgallate	12.31	473.108	M+H	0.2	455, 289, 167, 139
27	Epigallocatechin 3-coumaroate	16.23	453.1178	M+H	-0.4	435, 209, 139
28	Epiafzelechin*	13.39	275.0912	M+H	-0.7	191, 139, 107, 55

* Confirmed by standard compounds

Table 3. List of identified compounds with retention time, accurate mass, mass error and their MS2 fragment (continued).

No.	Identified Compound	RT (min)	Accurate Mass (*m/z*)	Adduct	Error (ppm)	MS2 Fragments
29	Epiafzelechin 3-gallate	15.92	427.1027	M+H	0.7	275, 153, 139, 107
30	1-Methylxanthine	3.97	167.0563	M+H	-0.6	110, 55
31	Adenosine*	3	268.1042	M+H	0.8	136
32	Guanosine*	3.31	284.0992	M+H	0.7	152, 135
33	5′-Methylthioadenosine	8.22	298.0968	M+H	-0.3	136
34	(S)-5′-Deoxy-5′-(methylsulfinyl)adenosine	2.58	314.0919	M+H	0.3	164, 136, 97
35	Cytidine 2′-phosphate	1.73	324.0595	M+H	1.2	112
36	Cyclic AMP	2.35	330.0594	M+H	-1.2	312, 136
37	AMP*	1.87	348.0707	M+H	0.9	136
38	ADP*	1.77	428.0362	M+H	-1.2	348, 136
39	Caffeoylshikimic acid	12.64	337.0891	M+H	-8	163, 145, 135, 117
40	3-*p*-Coumaroylquinic acid	11.95	339.1075	M+H	0	165, 147, 119
41	4-*p*-Coumaroylquinic acid	8.89	339.1074	M+H	-0.3	147, 119
42	1-Caffeoylquinic acid	11.75	355.1021	M+H	-0.8	337, 309, 163, 121
43	4-Caffeoylquinic acid	7.13	355.1021	M+H	-0.8	337, 309, 163, 121
44	3,5-Di-caffeoylquinic acid	17.26	517.1337	M+H	-0.8	471, 453, 163, 121
45	Chlorogenic acid*	9.71	355.1025	M+H	0.3	163, 145, 117, 89
46	Theogallin*	3.76	345.0817	M+H	0.3	327, 171, 153, 125
47	Quinic acid*	1.72	193.0711	M+H	2.1	149, 111, 95, 83
48	Strictinin*	8.1	652.1146	M+NH4	0.3	482, 447, 303, 277
49	Trigalloyl glucose	12.12	654.13	M+NH4	-0.1	467, 297, 153
50	Benzyl primeveroside	11.81	425.142	M+Na	0.5	331, 255, 179, 153
51	Phenylethyl primeveroside	14.46	455.1294	M+K	-4.4	351
52	Linalool primeveroside	25.99	471.2205	M+Na	1	335, 333
53	Linalool oxide primeveroside	20.5	487.2156	M+Na	1.3	335, 333
54	Linalool oxide primeveroside isomer	18.08	503.191	M+K	-0.4	351
55	Proanthocyanidin A1	3.44	577.1334	M+H	-1.2	559, 425, 407, 121
56	Procyanidin B1*	8.39	579.1501	M+H	0.7	409, 291, 289, 127
57	Procyanidin B2*	9.46	579.1503	M+H	1	409, 301, 289, 127
58	Procyanidin B3	7.12	579.1496	M+H	-0.2	409, 271, 127
59	Procyanidin B5	7.41	579.1494	M+H	-0.5	427, 409, 289, 127
60	Procyanidin C1	8.21	867.2133	M+H	0.2	849, 591, 153, 139
61	Theaflavin*	24.29	565.134	M+H	-0.2	427, 259, 139
62	Theaflavin-3-gallate*	24.05	717.1447	M+H	-0.4	699, 397, 153, 139
63	Theaflavin 3,3′-digallate*	24.28	869.1554	M+H	-0.7	731, 561, 333, 139
64	Theasinensin A	8.53	915.1617	M+H	0.2	897, 763, 139, 153
65	Theasinensin B	5.9	763.1514	M+H	1.2	595, 443, 305, 139
66	Theasinensin C	4.82	611.1392	M+H	-1.2	593, 139
67	Theasinensin F	9.82	899.167	M+H	-0.9	425, 287, 153
68	Luteolin-8-*C*-glucoside*	14.75	449.1077	M+H	-0.5	353, 329, 299

* Confirmed by standard compounds

Table 3. List of identified compounds with retention time, accurate mass, mass error and their MS2 fragment (continued).

No.	Identified Compound	RT (min)	Accurate Mass (*m/z*)	Adduct	Error (ppm)	MS2 Fragments
69	Vitexin*	17.78	433.1126	M+H	-0.9	313, 139, 85
70	Isovitexin*	18.13	433.1126	M+H	-0.9	313, 283, 121, 81
71	Isovitexin 2″-*O*-glucoside	11.55	595.1654	M+H	-0.6	433, 313, 139, 85
72	Quercetin 3-*O*-galactoside*	18.53	465.1031	M+H	0.7	303, 165, 91
73	Quercetin 3-*O*-glucoside*	18.88	465.1032	M+H	0.9	303
74	Quercetin 3-*O*-glucuronide	17.27	479.0815	M+H	-1.5	303
75	Rutin*	18.87	611.1613	M+H	1	465, 303, 85
76	Quercetin diglucoside	15.78	627.1554	M+H	-0.3	303
77	Quercetin 3-*O*-glucosylrutinoside*	17.75	773.2143	M+H	1	611, 465, 303
78	Quercetin 3-*O*-galactosylrutinoside*	17.59	773.2145	M+H	1.3	611, 465, 303, 145
79	Quercetin triglucoside	14.86	789.2085	M+H	0.1	303
80	Myricitrin*	17.76	465.103	M+H	0.4	319
81	Myricetin 3-*O*-glucoside	15.78	481.098	M+H	0.6	319
82	Myricetin 3-*O*-galactoside*	15.61	481.0981	M+H	0.8	319, 127, 85
83	Apigenin-6,8-*C*-diglucoside*	12.69	595.1665	M+H	1.2	559, 475, 307, 153
84	Apigenin-6-*C*-glucosyl-8-*C*-arabinoside	14.54	565.1559	M+H	1.2	427, 409, 391, 379
85	Apigenin-6-*C*-arabinoside-8-*C*-glucoside*	15	565.1559	M+H	1.2	547, 529, 511, 469
86	Kaempferol 3,7-dirhamnoside	16.7	579.1706	M+H	-0.4	287
87	Kaempferol 3-*O*-arabinoside*	21.59	419.0971	M+H	-0.5	287
88	Kaempferol 3-*O*-glucoside*	21.23	449.1082	M+H	0.7	287, 85
89	Kaempferol 3-*O*-galactoside*	20.58	449.1083	M+H	0.9	287
90	Kaempferol 7-*O*-rutinoside	20.59	595.1656	M+H	-0.3	449, 287
91	Kaempferol 3-*O*-rutinoside*	21.31	595.1663	M+H	0.8	449, 287, 147, 331
92	Kaempferol 7-(6″-galloylglucoside)	18.83	601.1186	M+H	-0.3	287, 153, 125
93	Kaempferol 3-(6″-galloylglucoside)	19.94	601.1186	M+H	-0.3	287, 153, 125
94	Myricetin 3-(3″-galloylrhamnoside)	8.98	617.1134	M+H	-0.5	319, 153, 125
95	Kaempferol 3-*O*-glucosylrutinoside*	20.37	757.2192	M+H	0.8	595, 449, 287, 331
96	Kaempferol 3-*O*-galactosylrutinoside*	19.49	757.2197	M+H	1.5	595, 449, 287, 331
97	N-(1-Deoxy-1-fructosyl)leucine	3.41	294.1551	M+H	1	276, 258, 230, 212
98	N-(1-Deoxy-1-fructosyl)phenylalanine	5.27	328.1392	M+H	0.3	292, 264, 132, 120
99	Theanine glucoside	2.22	337.1605	M+H	-0.3	301, 208, 158

* Confirmed by standard compounds

Metabolite variation among the subtypes of white tea

Hierarchy cluster analysis was used to visualize the abundance variations for the 64 out of 99 identified metabolites with significant differences across the three subtypes of white tea (P ≤0.05). Figure 5 shows that each subtype of white tea can accurately be grouped based on the patterns of the identified differential metabolites. These metabolites include primarily four groups of compounds:

- Catechins/dimeric catechins
- Phenolic acids/hydrolysable tannins/amino acids
- Flavonol glycosides/flavone glycosides
- Alkaloids/nucleotides/aroma precursors

Among the four groups of compounds, most flavonol glycosides/flavone glycosides show the lowest abundance in Silver Needle, relatively high abundance in White Poeny, and the highest abundance in Shoumei.

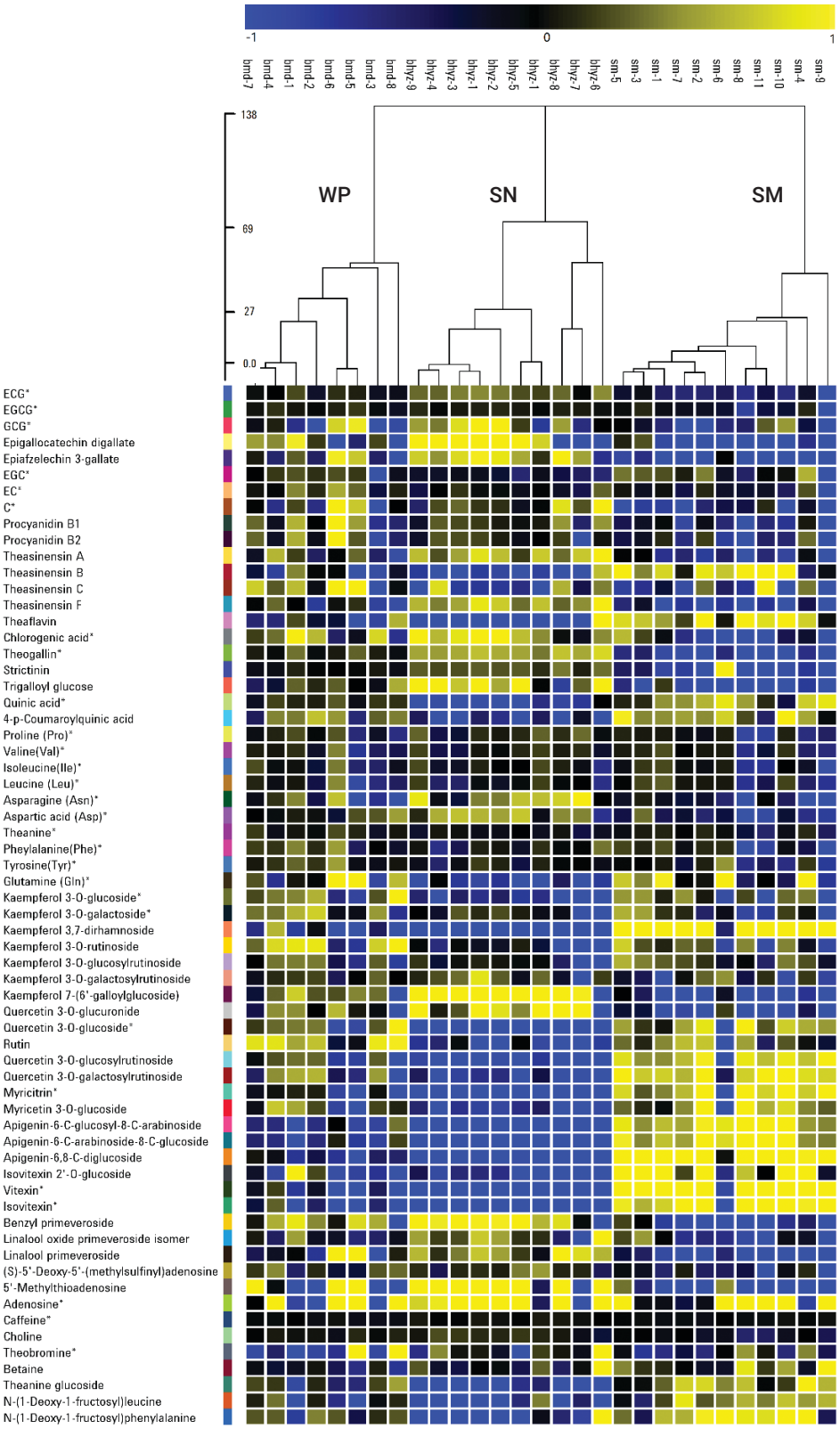

Figure 5. Hierarchy cluster analysis demonstrating the abundance variation of the significant metabolites among the three subtypes of white tea.

Correlation of the differential metabolite levels with tea flavor

Pearson correlation analysis was conducted to discover the metabolites contributing to the specific tea flavor including aspects such as umami, bitterness, and astringency. Figure 6A shows that 41 of 64 metabolites exhibit high correlation (R^2 ≥0.9) with either flavor. The yellow or blue color of the cell indicates that the compound abundance is either positively or negatively correlated, respectively, with the corresponding flavor. Figure 6A shows that most differential compounds exhibit consistent tendency for the three types of characteristic tea flavor. Figures 6B to 6D show the abundance variations for the major class of differential metabolites (P ≤0.01) in three subtypes of white tea.

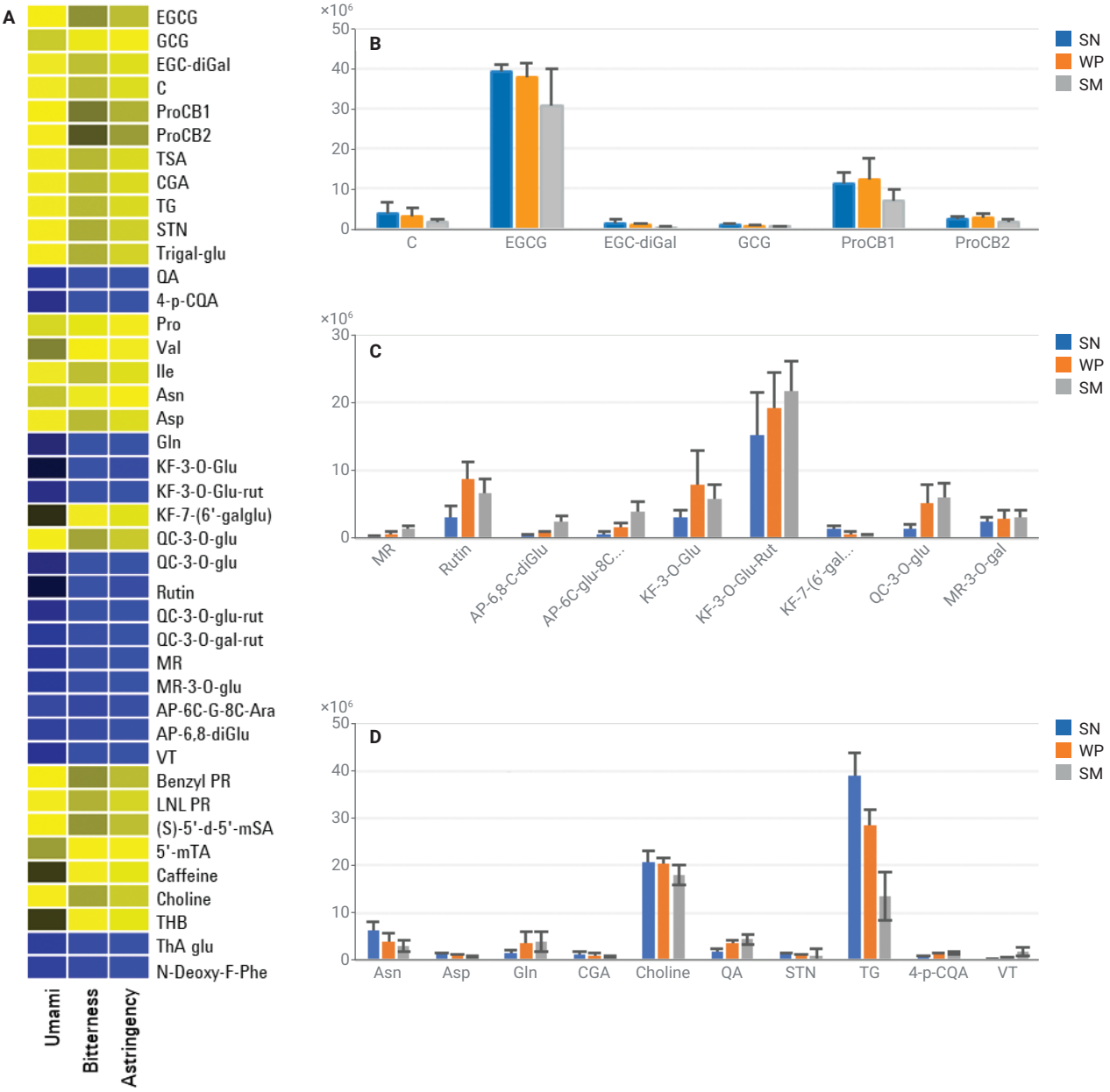

Figure 6. Correlation of the identified differential compound abundances with white tea quality in terms of umami, bitterness, and astringency taste. (A) Pearson correlation plot showing how differential compound abundances either positively (yellow) or negatively (blue) correlate with the three types of tea taste; (B to D) Variation of the abundance for the major differential compounds including catechins and dimeric catechins (B), flavonol glycosides and flavone glycosides (C), and phenolic acids, hydrolyzable tannins, and amino acids (D).

Catechins and dimeric catechins that positively correlate with the white tea flavor display a relatively higher presence in SN and WP than SM subtypes of white tea (Figure 6B). Flavonol glycosides and flavone glycosides negatively correlate with tea flavor, and most of them show a low presence in SN, and are relatively abundant in WP and SM (Figure 6C). Some acids such as Asn, TG, and choline showed a high presence in SN and WP, but a reverse tendency was observed for other acids (Figure 6D).

Conclusion

Nontargeted metabolomic profiling analysis was successfully performed for studies on white tea characteristics with three major subtypes. Among the identified differential metabolites, 64 showed significant abundance differences among subtypes of white teas, and 41 exhibited excellent correlation with the umami, bitterness, and astringency of the white teas. This indicates that these metabolites contribute primarily to the flavor characteristics of the subtypes of tea, and have the potential to serve as markers for quality control of white tea.

References

1. Mao, J. T.; *et al.* White Tea Extract Induces Apoptosis in Non–Small Cell Lung Cancer Cells: the Role of Peroxisome Proliferator-Activated Receptor-γ and 15-Lipoxygenases. *Cancer Prev. Res.* **2010**, *3(9)*, 1132–1140.
2. Ning, J-M.; *et al.* Chemical Constituents Analysis of White Tea of Different Qualities and Different Storage Times. *Eu. Food Res. Technol*. **2016**, *242(12)*, 2093–2104.
3. Tan, J.; *et al.* Flavonoids, Phenolic Acids, Alkaloids and Theanine in Different Types of Authentic Chinese White Tea Samples. *J. Food Compos. Anal.* **2017**, *57*, 8–15.
4. Dai, W.; *et al.* Nontargeted analysis using ultraperformance liquid chromatography-quadrupole time-of-flight mass spectrometry uncovers the effects of harvest season on the metabolites and taste quality of tea (*Camellia sinensis* L.). *J. Agric. Food Chem*. **2015**, *63*, 9869–9878.
5. Tan, J.; *et al.* Study of the dynamic changes in the non-volatile chemical constituents of black tea during fermentation processing by a non-targeted metabolomics approach. *Food Res. Intern*. **2016**, *79*, 106–113.

www.agilent.com/chem

This information is subject to change without notice.

© Agilent Technologies, Inc. 2019

Printed in the USA, April 9, 2019
5994-0866EN

Agilent
Trusted Answers