

# Characteristics of non-volatile metabolites in fresh shoots from tea plant (*Camellia sinensis*) and its closely related species and varieties

Chenkai Jiang<sup>1,2</sup>, Doo-Gyung Moon<sup>3</sup>, Jianqiang Ma<sup>1\*</sup>, and Liang Chen<sup>1\*</sup>

<sup>1</sup> Key Laboratory of Biology, Genetics and Breeding of Special Economic Animals and Plants, Ministry of Agriculture and Rural Affairs; Tea Research Institute of the Chinese Academy of Agricultural Sciences, Hangzhou 310008, China

<sup>2</sup> Institute of Sericulture and Tea, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China

<sup>3</sup> Research Institute of Climate Change and Agriculture, National Institute of Horticultural and Herbal Science, Jeju 690-150, Korea

\* Corresponding authors, E-mail: [majianqiang@tricaas.com](mailto:majianqiang@tricaas.com); [liangchen@tricaas.com](mailto:liangchen@tricaas.com)

## Abstract

Tea plant (*Camellia sinensis*) and its closely related species and varieties belong to Sect. *Thea* (L.) Dyer, *Camellia* L. There are abundant compounds in the fresh shoots of section *Thea* (L.) Dyer species and varieties. Their variation in different tea species and varieties is unclear. Fresh shoots from 336 accessions of *C. sinensis* and its closely related species and varieties were harvested and their non-volatile metabolites were detected through UPLC-MS (ultra-performance liquid chromatography - mass spectrometry). A total of 374 non-volatile metabolites were identified, which can be divided into 27 categories. Among them, 32 compounds were flavonoid polymers. The tea plants were divided into two groups, according to the Calinski criterion according to the composition of metabolites. The top 30 differential metabolites in *C. sinensis* var. *sinensis*, *C. sinensis* var. *assamica*, *C. sinensis* var. *pubilimba*, *C. tachangensis*, and *C. taliensis*, belong to amino acids and their derivatives, benzoic acid derivatives, carbohydrates, coumarins, flavonol glycosides, organic acids, quinoline acid and its derivatives. The results provide new insights for further understanding the characteristic metabolites of tea plant and its closely related species and varieties.

**Citation:** Jiang C, Moon DG, Ma J, Chen L. 2022. Characteristics of non-volatile metabolites in fresh shoots from tea plant (*Camellia sinensis*) and its closely related species and varieties. *Beverage Plant Research* 2:9 <https://doi.org/10.48130/BPR-2022-0009>

## INTRODUCTION

The tea plant is a perennial and economic crop with a productive lifespan. Its fresh shoots are processed into tea, including green, black, Oolong, white, yellow and dark tea<sup>[1]</sup>. Numerous metabolites in fresh tea shoots are precursors which experience a series of variations during the tea manufacturing process and contribute significantly to tea quality and flavor<sup>[2]</sup>. Additionally, these metabolites play important roles in resistance for tea itself and functional effects on human health. Flavonoids, theanine, caffeine and theobromine are the dominant metabolites in section *Thea* (L.) Dyer<sup>[1]</sup>.

Tea is basically produced from three kinds of tea plants: *Camellia sinensis* (L.) O. Kuntze var. *sinensis*, *C. sinensis* var. *assamica* (Masters) Kitamura and *C. sinensis* var. *pubilimba* Chang. As well as these, wild tea plants, involving *C. tachangensis* F. C. Zhang, *C. taliensis* (W. W. Smith) Melchior, *C. crassicolumna* Chang, *C. gymnogyna* Chang also belong to sect. *Thea* (L.) Dyer. The sect. *Thea* is divided into the above mentioned species and varieties depending on the ovary chamber number, the style splitting number and the presence or absence of ovary pubescence, combined with the corolla size and the morphology of tree, branch and leaves<sup>[3]</sup>. *C. tachangensis* was first identified in Yunnan, China in the 1980's<sup>[2]</sup>. *C. taliensis* is usually distributed in the evergreen broad-leaved forests from 1,300 to 2,400 m above sea level, mainly growing in southwestern Yunnan in China and its adjacent regions<sup>[2]</sup>. There are significant differences among *C. tachangensis*, *C. taliensis*, *C. gymnogyna* and *C. sinensis* based on

the chloroplast genome<sup>[2]</sup>. A number of papers have presented the metabolites in leaves from *C. sinensis* with the rapid development of metabolomics technology. They reveal the effects of temperature, light, water, fertilizer, climate, altitude and other factors on metabolites of the tea plant.

Some studies have researched the metabolic characteristic of wild tea plants. *C. crassicolumna* which has an abundance of phenols and no caffeine or theophylline in the leaves<sup>[2]</sup>. Zhang et al.<sup>[2]</sup> found that *C. tachangensis* had significantly higher contents of epicatechin and epigallocatechin compared to *C. sinensis*. In contrast, Yu<sup>[4]</sup> believed that *C. tachangensis* had low accumulation of polyphenols, especially low content of catechins. Gujing tea (*C. gymnogyna*) in Guizhou province has high content of proanthocyanidins and galloylquinic acid<sup>[5]</sup>, and this species in Guangxi province contains three different purine alkaloids, among which theobromine content was the highest<sup>[6]</sup>. However, there were less than three wild tea plants in each study, some conclusions are contradictory, and it is difficult to reflect their comprehensive metabolic profiles, let alone the comparison of metabolites among tea and its wild relatives. Therefore, we gathered fresh shoots from 336 accessions cultivated tea and its wild species and varieties in spring and detected their metabolic components and content. We compare their distinction of metabolites and aim to explore their characteristic metabolites. The current study elucidates the common and unique character of different tea plant species and varieties, especially flavonoids, amino acids and alkaloids.

## RESULTS

### Overview of metabolites in tea plants

The correlation between every two QC samples was calculated according to Pearson correlation coefficient. [Supplemental Fig. S1](#) shows correlation ranges from 0.92 to 1.00, suggesting that the instrument is stable. Three methods were performed to identify compounds: (1) Retention time, primary and secondary fragments of compounds were aligned to standards; (2) Primary and secondary fragments of compounds were aligned to the literature; (3) The remaining compounds were annotated through information from online databases. A total of 4,836 metabolites were detected, 374 of which were identified ([Supplemental Table S1](#)). Concretely, 117 and 18 of them were identified through standards and literature, respectively. The remaining 239 were identified via online databases.

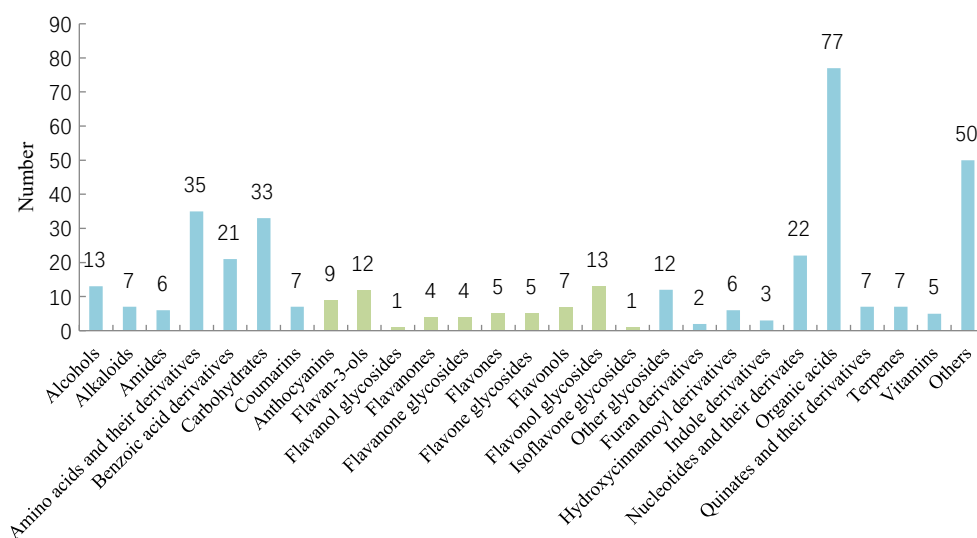
The identified metabolites were divided into 27 categories as presented in [Fig. 1](#), including alcohols, alkaloids, amides, amino acids and their derivatives, benzoic acid derivatives, carbohydrates, coumarins, anthocyanins, flavan-3-ols, flavanol glycosides, flavanones, flavanone glycosides, flavones, flavonols, flavonol glycosides, isoflavone glycosides, other glycosides, furan derivatives, hydroxycinnamoyl derivatives, indole derivatives, nucleotides and their derivatives, organic acids, quinic acids and their derivatives, terpenes, vitamins and others.

derivatives, hydroxycinnamoyl derivatives, nucleotides and their derivatives, organic acids, quinic acids and their derivatives, terpenes, vitamins and others.

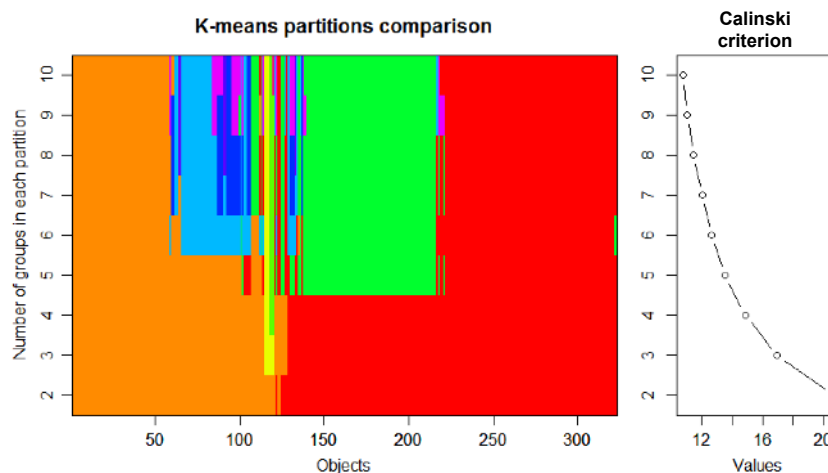
### Cluster analysis relying on the composition and content of metabolites

The tea accessions could be classified as CSS [*C. sinensis* (L.) O. Kuntze var. *sinensis*], CSA [*C. sinensis* var. *assamica* (Masters) Kitamura], CSP [*C. sinensis* var. *pubilimba* Chang], CTF (*C. tachangensis* F. C. Zhang), CTM [*C. taliensis* (W.W. Smith) Melchior], CCC (*C. crassicolumna* Chang) in Sect. *Thea* (L.) Dyer, and CRL (*C. reticulata* Lindl.), CCT (*C. crapnelliana* Tutch.), CMC (*C. muricatula* Chang) and COC (*C. oleifera* Abel.) are non-tea *Camellia* species. K value (Calinski criterion) was 2 based on composition and content of metabolites ([Fig. 2](#)). Therefore, tea samples were divided into two groups. Group 1 contained 301 tea accessions ([Supplemental Fig. S2](#)), including CSS, CSA, CSP, CTF, CTM and CPC. Group 2 contained 22 tea plants, including CSS, CSA, CSP, CRL, CCT and COC.

Different accumulated metabolites (DAMs) were determined after *t* test, whose fold change > 2, FDR < 0.05. The volcano plots of DAMs ([Fig. 3a](#)) reveal that the content of caffeine, astilbin, theacrine, 7-ethoxycoumarin, (2Z)-3-[4,5-dihydroxy-2-



**Fig. 1** Classification of metabolites for tea plant and its related species and varieties.



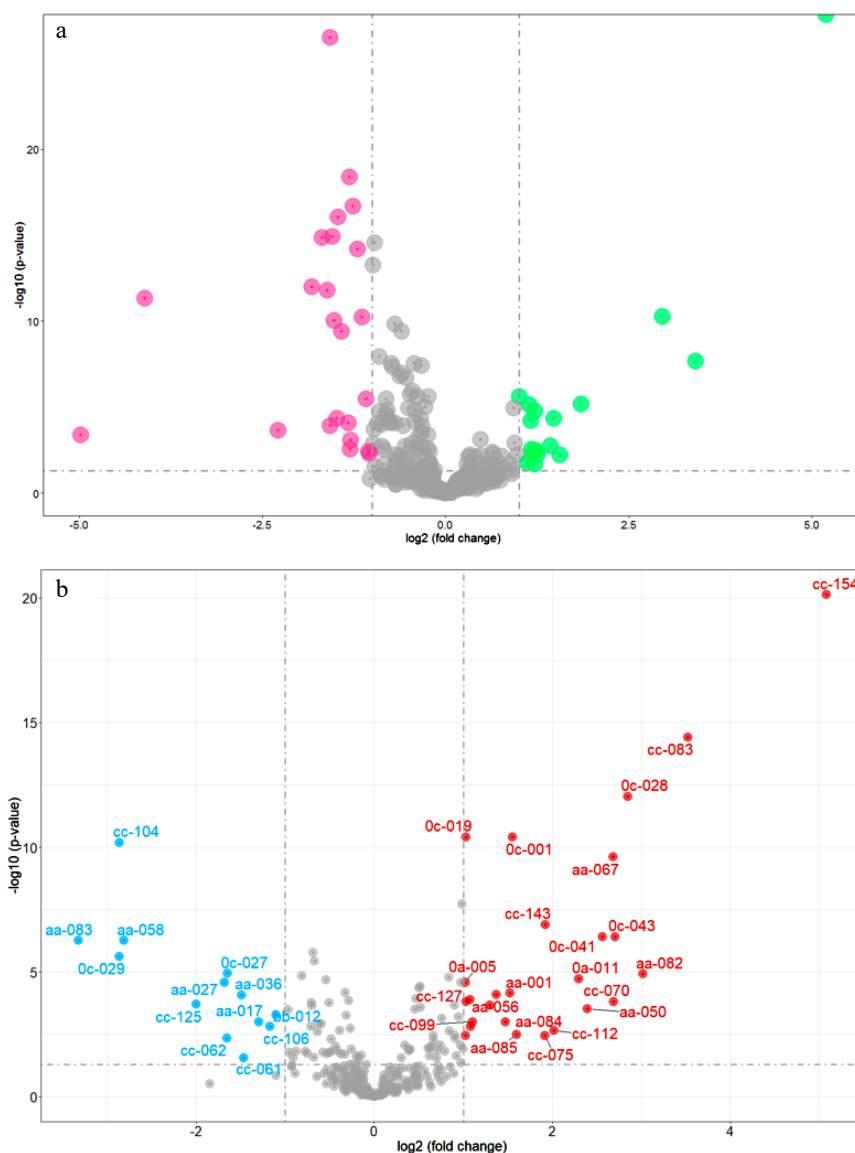
**Fig. 2** Group number of tea plants determined based on Calinski criterion.

## Metabolites of tea plants

(2-hydroxy-2-propanyl)-2,3-dihydro-1-benzofuran-7-yl]acrylic acid, olsalazine, triga-glucose, 3-[(1-carboxyvinyl)oxy] benzoic acid, 1,6-bis-O-galloyl- $\beta$ -D-glucose, matricin, malonylglycitin, GC-diGA, N-acetyl-DL-tryptophan, vitexin-2''-O-rhamnoside, glucosyl-galactosyl hydroxylysine, 3-hydroxy-10-(4-hydroxy-3-methoxyphenyl)-8-(3,5,7-trihydroxy-4-oxo-3,4-dihydro-2h-chromen-2-yl)-4-oxatricyclo[4.3.1.0~3,7~] decan-2-one, propentofylline and luteolin 7-O-(6-O-malonyl- $\beta$ -D-glucoside) was significantly higher in Group 1 than those in Group 2. While the content of procyanidin B1, 3,4-dihydroxybenzaldehyde, catechin, 4-hydroxybenzaldehyde, 1,4-anhydro-6-O-dodecanoyl-2,3-bis-O-(2-hydroxyethyl)-glucitol, sinapinaldehyde, procyanidin B4, 1-naphthyl glucuronide, gly-lys, 1,5-anhydro-2-O-(6-O-benzoyl- $\alpha$ -l-galactopyranosyl)-D-glucitol, theobromine, (2R,3R)-5,7-dihydroxy-2-(4-hydroxyphenyl)-3,4-dihydro-2h-chromen-3-yl, 3,4,5-trihydroxybenzoate, 2-hydroxycinnamic acid, 1-O-vanilloyl- $\beta$ -D-glucose, 3,22-dihydroxy-28-oxo-olean-

12-en-16-yl acetate, gallotanin 1,2,6-tri-O-galloyl- $\beta$ -D-glucopyranose, 1,2,3,6-tetra-O-galloyl- $\beta$ -D-glucose, 1,2-di-O-galloyl-HHDP-glucose, GC-GCG, agmatine, iso-quercitrin 6''-acetate, 2,3-dihydro-1,4-benzodioxine-5-carboxylic acid, zizybeoside I and 5-(2-carboxyethyl)-2-hydroxyphenyl  $\beta$ -D-glucopyranosiduronic acid was significantly lower in Group 1 than those in Group 2.

The content of four flavan-3-ols (ECG4''Me, EGCG4''Me, GC-diGA, GC-GCG), four organic acids (2,3-dihydro-1-benzofuran-2-carboxylic acid, ursolic acid, 3-acetyl-11-keto- $\beta$ -boswellic acid), three amino acids and their derivatives (L-glutathione, L-histidine L-arginine), 1 coumarin (7-hydroxycoumarine), two alkaloids (propentofylline, theacrine) is significantly higher in *C. sinensis* than its relatives (Fig. 3b). While the content of two quinate and its derivatives (neochlorogenic acid, chlorogenic acid), two flavonol glycosides (nictoflorin, iso-quercitrin 6''-acetate), three organic acids (1,2-di-O-galloyl-HHDP-glucose,



**Fig. 3** Volcano plots of DAMs. (a) DAMs of natural population (red dots represent the content of DAMs that are higher in Group 2 compared to Group 1, green dots represent the content of DAMs that are higher in Group 1 compared to Group 2); (b) DAMs between *Camellia sinensis* and its relatives (red dots represent the content of DAMs that are higher in *Camellia sinensis* compared to relatives of *Camellia sinensis*, blue dots represent the content of DAMs that are higher in relatives of *Camellia sinensis* compared to *Camellia sinensis*).

caffeic acid, phenylacetic acid) is significantly lower in *Camellia sinensis* than its relatives. Furthermore, 7-hydroxycoumarin, propentofylline, 1,2-di-O-galloyl-HHDP-glucose and theacrine are the DAMs whose fold change  $\geq 3$  (Fig. 3b).

### Metabolomic characteristic of different species and varieties of Sect. *Thea*

CRL, CCC, CMC, COC were deleted before DAM analysis among Sect. *Thea* plants as their sample number was less than 4. The remainder of the tea plants were CSS, CSA, CSP, CTF, CTM. The top 30 DAMs are displayed in Fig. 4. A total of 16 metabolites were in the top 30 in two consecutive years, involving amino acids and their derivatives, benzoic acid derivatives, carbohydrates, coumarins, quinate and their derivatives

(Table 1). These metabolites could be used to discriminate different Sect. *Thea* plants. Among them, the content of L-pyroglutamic acid, L-serine, sinapaldehyde glucoside, 7-ehoxycoumarin and D - quinic acid was highest in CSS. The content of N-acetyl-DL-tryptophan and azelaic acid was highest in CSA. The content of 7-hydroxycoumarin, diGC-GA, 2-acetamido-2-deoxyglucose, kaempferitrin, nictoflorin, traumatic acid and neochlorogenic acid was highest in CTF. The content of 5'-xanthylic acid and chlorogenic acid was highest in CTM.

### Polymers of flavonoids

There are abundant polymers of flavonoids in the identified metabolites (Table 2). A total of 32 flavonoid polymers were

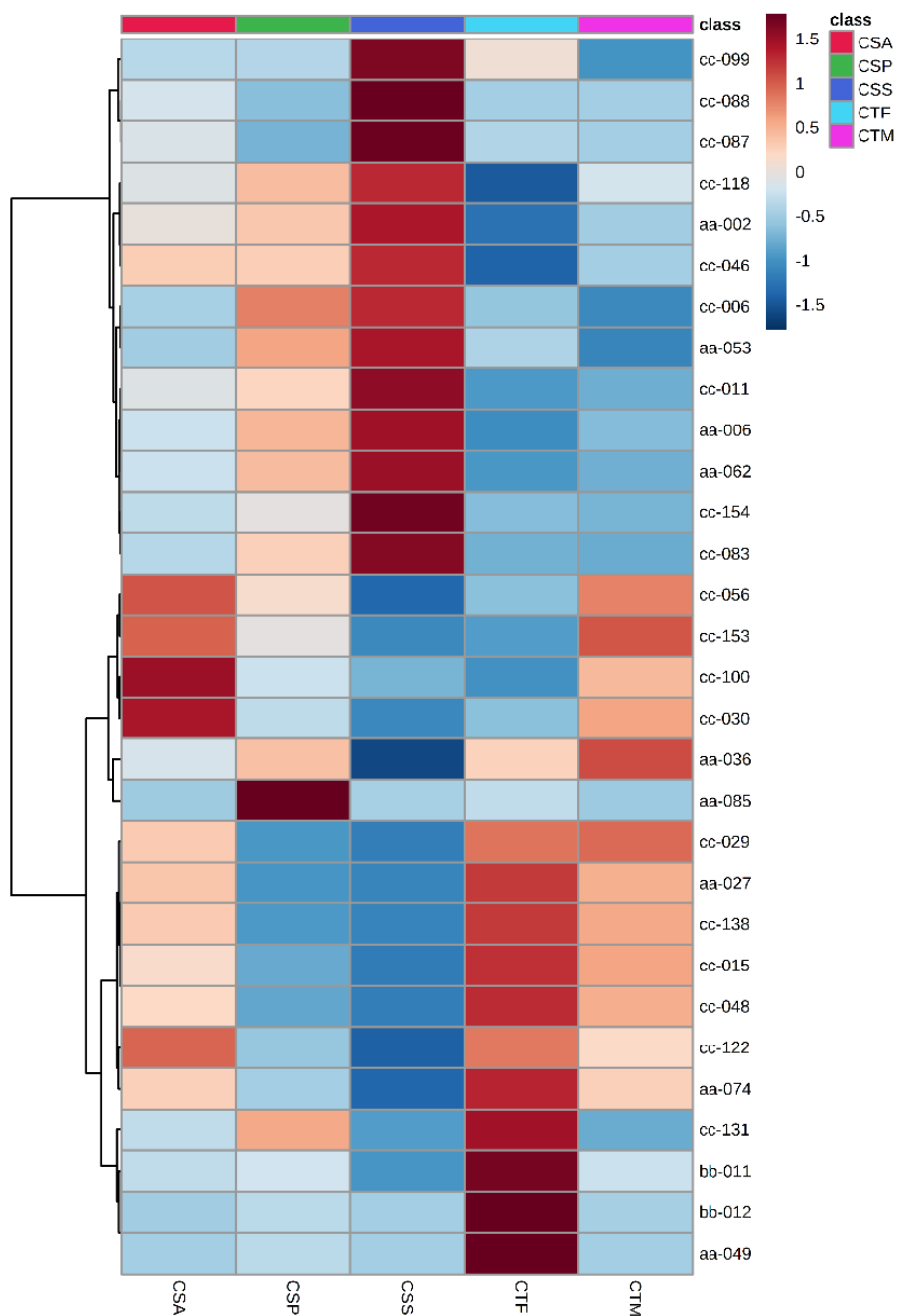


Fig. 4 Top 30 DAMs of five tea plant species and varieties.

**Table 1.** Signature metabolites of five tea plant species and varieties.

Representative group*	Metabolite	Class	Relative peak area				
			CSA	CSP	CSS	CTF	CTM
CSS	L-Pyroglutamic acid	Amino acids and their derivatives	5.E+07	7.E+07	8.E+07	4.E+07	4.E+07
CSS	L-Serine	Amino acids and their derivatives	1.E+06	1.E+06	2.E+06	1.E+06	1.E+06
CSS	Sinapaldehyde glucoside	Carbohydrates	3.E+05	2.E+05	4.E+05	2.E+05	2.E+05
CSS	7-Ethoxycoumarin	Coumarins	1.E+05	8.E+04	2.E+05	1.E+05	8.E+04
CSS	D-Quinic acid	Quinates and their derivatives	3.E+06	3.E+06	4.E+06	2.E+06	2.E+06
CSA	N-Acetyl-DL-tryptophan	Amino acids and their derivatives	1.E+06	4.E+05	3.E+05	3.E+05	1.E+06
CSA	Azelaic acid	Organic acids	2.E+05	1.E+05	9.E+04	2.E+05	2.E+05
CTF	7-Hydroxycoumarine	Coumarins	3.E+06	2.E+06	3.E+06	4.E+06	3.E+06
CTF	diGC-GA	Benzoic acid derivatives	2.E+07	2.E+07	1.E+07	3.E+07	2.E+07
CTF	2-Acetamido-2-deoxyglucose	Carbohydrates	2.E+06	2.E+06	1.E+06	3.E+06	2.E+06
CTF	Kaempferitrin	Flavonol glycosides	2.E+04	5.E+05	2.E+04	6.E+06	3.E+04
CTF	Nictoflorin	Flavonol glycosides	2.E+05	7.E+05	2.E+05	7.E+06	2.E+05
CTF	Traumatic acid	Organic acids	5.E+05	3.E+05	3.E+05	6.E+05	5.E+05
CTF	Neochlorogenic acid	Quinates and their derivatives	7.E+06	2.E+06	2.E+06	1.E+07	9.E+06
CTM	5'-Xanthylic acid	Quinates and their derivatives	5.E+05	3.E+05	3.E+05	5.E+05	6.E+05
CTM	Chlorogenic acid	Quinates and their derivatives	1.E+07	1.E+07	4.E+06	9.E+06	2.E+07

\* CSS: *C. sinensis* var. *sinensis* (L.) O. Kuntze; CSA: *C. sinensis* var. *assamica* (Masters) Kitamura; CSP: *C. sinensis* var. *pubilimba* Chang; CTF: *C. tachangensis* F. C. Zhang; CTM: *C. taliensis* (W. W. Smith) Melchior.

**Table 2.** Polymers of flavonoids.

Name	Classification	Formula
GC-GCG	Anthocyanins	C <sub>37</sub> H <sub>30</sub> O <sub>18</sub>
Procyanidin B1	Anthocyanins	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>
Procyanidin B3	Anthocyanins	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>
Procyanidin B4	Anthocyanins	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>
Procyanidin C1	Anthocyanins	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>
Tricatechins1	Anthocyanins	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>
Tricatechins2	Anthocyanins	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>
Tricatechins3	Anthocyanins	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>
GC-diGA	Benzoic acid derivatives	C <sub>29</sub> H <sub>22</sub> O <sub>15</sub>
Eriocitrin	Flavanol glycosides	C <sub>27</sub> H <sub>32</sub> O <sub>15</sub>
Eriodictyol C-hexoside	Flavanone glycosides	C <sub>21</sub> H <sub>22</sub> O <sub>11</sub>
Naringin	Flavanone glycosides	C <sub>27</sub> H <sub>32</sub> O <sub>14</sub>
Naringin dihydrochalcone	Flavanone glycosides	C <sub>27</sub> H <sub>34</sub> O <sub>14</sub>
Prunin	Flavanone glycosides	C <sub>21</sub> H <sub>22</sub> O <sub>10</sub>
Baicalin	Flavone glycosides	C <sub>21</sub> H <sub>18</sub> O <sub>11</sub>
Cynaroside	Flavone glycosides	C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>
Luteolin 7-O-(6-O-malonyl-β-D-glucoside)	Flavone glycosides	C <sub>24</sub> H <sub>22</sub> O <sub>14</sub>
Schaftoside	Flavone glycosides	C <sub>26</sub> H <sub>28</sub> O <sub>14</sub>
Astilbin	Flavonol glycosides	C <sub>21</sub> H <sub>22</sub> O <sub>11</sub>
Baimaside	Flavonol glycosides	C <sub>27</sub> H <sub>30</sub> O <sub>17</sub>
Isoquercetin	Flavonol glycosides	C <sub>21</sub> H <sub>20</sub> O <sub>12</sub>
Kaempferin	Flavonol glycosides	C <sub>21</sub> H <sub>20</sub> O <sub>10</sub>
Kaempferitrin	Flavonol glycosides	C <sub>27</sub> H <sub>30</sub> O <sub>14</sub>
Myricetin 3-O-galactoside	Flavonol glycosides	C <sub>21</sub> H <sub>20</sub> O <sub>13</sub>
Nictoflorin	Flavonol glycosides	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>
Quercetin 3-O-α-D-xylopyranoside	Flavonol glycosides	C <sub>20</sub> H <sub>18</sub> O <sub>11</sub>
Quercitrin	Flavonol glycosides	C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>
Rutin	Flavonol glycosides	C <sub>27</sub> H <sub>30</sub> O <sub>16</sub>
Tilioside	Flavonol glycosides	C <sub>30</sub> H <sub>26</sub> O <sub>13</sub>
Vitexin	Flavonol glycosides	C <sub>21</sub> H <sub>20</sub> O <sub>10</sub>
Vitexin-2"-O-rhamnoside	Flavonol glycosides	C <sub>27</sub> H <sub>30</sub> O <sub>14</sub>
Malonylgenistin	Isoflavone glycosides	C <sub>24</sub> H <sub>22</sub> O <sub>13</sub>

identified, consisting of flavonoid, sugar and gallic acid. Most of them were catechin polymers and flavonoid glycosides. More interestingly, flavonol was the major monomer of flavonoid glycosides. A total of 23 flavonol glycosides were identified in

the current study, whose major glycoside ligands were glucose and rhamnose. In addition, almost all the glycoside ligands were hexose.

## DISCUSSION

The non-volatile compounds in tea could be divided into six groups, which were hydrolysable tannins, flavan-3-ols and their derivatives, flavonoids, alkaloids and theanine, simple phenols and their derivatives, terpenes<sup>[13]</sup>. In the present study, the non-volatile compounds in tea were divided into 27 groups, according to functional group integrated with chemical structure. Organic acids possessed the largest number among all the classifications. Gallic acid, benzoic acid, chlorogenic acid, α-ketoglutaric acid, malic acid, citric acid and oxalic acid were the predominant organic acids. Among them, organic acids play an important role in the metabolism of tea plants and tea quality usually as intermediate products from carbohydrate decomposition. The stability, water-solubility and activity of benzoic acid were changed after methylation or glycosylation<sup>[7]</sup>. Benzoic acid, chlorogenic acid, gallic acid and their derivatives were the essential phenolic acids in tea plants. Their content decreased with the maturity of tea leaves<sup>[8]</sup>. The content of benzoic acid near to the flowers declined during the flowering process<sup>[9]</sup>. The content of gallic acid was dramatically increasing in the process of green and black tea manufacturing<sup>[10]</sup>. White and green tea contained high-content chlorogenic acid, with white tea containing the most<sup>[11]</sup>. During deoxidation of Oolong tea, gallic acid was generated from the degradative ester-catechin<sup>[12]</sup>. These studies imply that organic acids are of great importance for the growth and development of tea plant and the quality of tea.

Tea plants in a narrow sense mean *C. sinensis* and its varieties. Tea plants in a broad sense, i.e., the Sect. *Thea* (L.) Dyer, include *C. tachangensis* F. C. Zhang, *C. crassicumna* Chang, *C. taliensis* (W.W. Smith) Melchior, *C. gymnogyna* Chang, *C. sinensis* (L.) O. Kuntze and its three varieties *C. sinensis* (L.) O. Kuntze var. *sinensis*, *C. sinensis* var. *assamica* (Masters) Kitamura and *C. sinensis* var. *pubilimba* Chang. The last two varieties are currently widely cultivated tea plants in China. *C. sinensis* (L.) O.

Kuntze var. *sinensis* is mainly distributed south of the Yangtze River. It has disseminated eastward to Japan and Korea, westward to Tibet, China and southward to northern Burma<sup>[13]</sup>. *C. sinensis* var. *assamica* (Masters) Kitamura is mainly distributed in Yunnan and its neighboring regions and countries, such as Vietnam, Thailand and Burma<sup>[13]</sup>. The majority of wild tea plants are CTF, CCC, CTM and CGC. CTF grows in eastern Yunnan, southwestern Guizhou and western Guangxi<sup>[14]</sup>. Compared to CSS, CTF contains higher amounts of EC, EGC, di-EC, ECG-EC, EGC-EC and tri-EC, with lower content of ester catechins<sup>[14]</sup>. There are abundant phenols in CCC. A total of 18 phenols were identified in *C. crassicaolumna* var. *multiplex*, involving four flavanols, six flavonol glycosides, three hydrolysable tannins, two derivatives from chlorogenic acid, and three simple phenols<sup>[15]</sup>. There was no caffeine or theophylline in one *C. crassicaolumna* var. *multiplex*<sup>[13]</sup>. In two CCC, the content of caffeine was 0.83% and 0.059%, the content of theobromine was 0.05% and 0.07%<sup>[16]</sup>. However, in the present study,  $3.26 \pm 0.09\%$  of caffeine was detected in one CCC through HPLC. It is likely that there are differences in alkaloid content among various CCC. Further CCC samples need to be gathered to explore their metabolic characteristics in the future. CTM, one of the most predominant tree varieties in southern tropical evergreen broad-leaved forest, is distributed from the middle of Lancang River in Hengduan Mountains of Yunnan to the Irrawaddy River Basin<sup>[17]</sup>. CTM is the relative of cultivated tea plant<sup>[18]</sup>. CTM from Lincang had nine phenols containing eight hydrolysable tannins, six catechin derivatives, three quinic acid aromatic esters, and two simple phenolics identified, along with caffeine<sup>[19]</sup>. Theobromine, caffeine and theacrine were detected in one CGC from Dayao Mountain and theobromine was its major alkaloid<sup>[6]</sup>. The content of theacrine was 3.2 times that of caffeine in CGC<sup>[20]</sup>. The results in the current study demonstrate that the composition and percentage of amino acids and their derivatives, benzoic acid derivatives, carbohydrates, coumarins, flavonol glycosides, nucleotide acids and their derivatives, organic acids, quininic acids and their derivatives could be distinct to different Sect. *Thea*. Flavonoids are the essential secondary metabolites in tea plant, closely related with flavor and the quality of brewed tea<sup>[21]</sup>. The convergence of soup is mainly caused by these substances<sup>[1]</sup>. Glycosylation is a ubiquitous modification process in plants<sup>[22]</sup>. Many substances have changed through glycosylation stability, solubility, biological activity and pharmacokinetics, which also helps tea plants resist biotic and abiotic stresses<sup>[22]</sup>. Glycosides are important metabolites in plants, which have the function of free radical scavenging, anti-radiation, anti-diabetes, and enhancing immunity<sup>[23]</sup>. Various flavonoid glycosides are formed in plants under the catalysis of UFGT. Phenolic acids, flavonols, flavonoids and isoflavones are the main substrates of UFGT<sup>[24]</sup>. Tea shoots are rich in flavonoid polymers, mainly flavonoid glycosides and anthocyanins. Natural flavonoids usually exist in glycoside form. The precursors of flavonols, flavonoids, anthocyanins, saponins and aroma substances mainly exist as glycosides in tea. The solubility of flavonoid glycosides is greater than that of flavonoids. Thus, flavonoid glycosides have greater influence on tea taste than flavonoids. The content of naringenin-7-O-glucoside, quercetin-3-O-glucoside and kaempfer-3-O-(6"-O-p-coumaryl)-glucoside in fresh tea leaves in August are higher than those in other seasons<sup>[25]</sup>. It could explain the reason why summer and autumn tea are

more bitter and astringent. Myrice-3-O-rhamnoside, myrice-3-O-galactoside, myrice-3-O-glucoside, quercetin-3-O-glucosyl-rhamnosyl-galactoside, quercetin-3-O-glucosyl-rhamnosyl-glucoside, vitexin-2"-O-rhamnoside, kaempfer-3-O-rutinoside, kaempfer-3-O-galactoside and kaempfer-3-O-glucoside were detected in tea<sup>[26]</sup>. A total of 202 glycosides in green tea were identified, and the sugar ligands were mainly glucose, galactose, rhamnose, rucose and primrose<sup>[27]</sup>. The higher the content of glycosides binding to EC, the lower the dimer content of EC, which indicates that the glycosides of EC are important donors of EC dimer<sup>[16]</sup>. It is likely that theanine converted into *N*-ethylpyrrolidone after deamination and decarboxylation. Then it may combine with the catechin A ring at 6-C or 8-C position to form theanine-catechin complex<sup>[28]</sup>. However, the polymers of theanine and catechin were not detected in this study, which may be due to the low content of theanine and catechin or the low response value of the instrument. It deserves establishing specific methods to detect these polymers in future.

## CONCLUSIONS

A total of 374 non-volatile metabolites were identified in tea plant (*Camellia sinensis*) and its closely related species and varieties. Organic acids and flavonoid are the most abundant and special non-volatile metabolites in tea plants, respectively. 7-hydroxycoumarine, propentofylline, 1,2-di-O-galloyl-HHDP-glucose and theacrine are the top four DAMs between *C. sinensis* and its relatives. Seven categories of substances: amino acids and their derivatives, benzoic acid derivatives, carbohydrates, coumarin, flavonol glycosides, organic acids, quinic acid and its derivatives are the top 30 DAMs among CSS, CSA, CSP, CTF, and CTM.

## MATERIALS AND METHODS

### Plant materials

The samples were collected from 336 tea accessions from Sect. *Thea* (L.) Dyer, *Camellia* L. and non-tea *Camellia* species, namely, 209 accessions of *C. sinensis* (L.) O. Kuntze var. *sinensis*, 66 accessions of *C. sinensis* var. *assamica* (Masters) Kitamura, 32 accessions of *C. sinensis* var. *pubilimba* Chang, four accessions of *C. tachangensis* F. C. Zhang, 15 accessions of *C. taliensis* (W. W. Smith) Melchior, one accession of *C. crassicaolumna* Chang, five accessions of *Camellia* sp., and one accession of *C. muricatula* Chang, *C. reticulata* Lindl., *C. crapnelliana* Tutch and *C. oleifera* Abel. (Supplemental Table S2). In the first round, two leaves and a bud from healthy tender shoots were harvested from March 16 to April 30, 2019. All the natural population tea plants were cultivated under the same horticultural conditions and management in the China National Germplasm Hangzhou Tea Repository in Hangzhou, China. Samples were fixed in liquid nitrogen immediately after plucking and were stored at  $-80^{\circ}\text{C}$ .

### Sample extraction

Ten milliliters of 70% methanol were added to  $0.2 (\pm 0.001)$  g of tea powder. The mixture was treated with ultrasonic extraction for 0.5 h. The supernatant was passed through a  $0.22 \mu\text{m}$  film after standing at  $4^{\circ}\text{C}$  for 2 h. The extracts were taken in a brown injection bottle and stored at  $-80^{\circ}\text{C}$  until injection.

## Metabolites of tea plants

Then, 100  $\mu\text{L}$  of the extract was taken as a Quality control (QC). And 10  $\mu\text{L}$  internal standard (0.025 mg/mL of sulfacetamide) was added to the sample before testing.

## LC-MS conditions

Metabolite detection was performed using a UPLC (Thermo Scientific Dionex Ultimate 3000, USA)-Q-Orbitrap (Thermo Scientific Q Exactive, USA) with a column C18 (SB-AQ, 1.8  $\mu\text{m}$ , 2.1 mm  $\times$  100 mm, Agilent, USA). Mobile phase A and B were 0.1% (v/v) formic acid in water and 0.1% (v/v) formic acid in acetonitrile. The gradient elution program was as follows: 0–6 min, 5%–20% B; 6–10 min, 20%–95% B; 10–11.5 min, 95% B; 11.5–15 min, 95%–5% B. The sample chamber temperature and the column temperature were 4  $^{\circ}\text{C}$  and 40  $^{\circ}\text{C}$ , respectively. The injection volume was 2  $\mu\text{L}$  with 0.3 mL/min. The relative collision energies were 15, 30, and 60. The m/z scanning range was 70–1,000. The spray voltage was 3.5 kV with the drying gas temperature at 320  $^{\circ}\text{C}$ , the auxiliary gas temperature at 350  $^{\circ}\text{C}$ , and the protective gas flow rate at 40 arb. The primary resolution was 70,000. The automatic gain control was 1e6. The secondary resolution was 17,500, and the automatic gain control was 1e5.

## Construction of the standard data base

Standards were dissolved with 70% methanol to 1.00 mg/mL as technical concentration. The mother liquid was diluted to 50 ng/mL.

## Metabolic data processing

Information such as substance name, mass-to-charge ratio, file location retrieved from standards and literature were input to mzValut to build a local library. The local library and sample information was imported into Compound discoverer 2.1 to characterize and quantify metabolites. The substances in the local library were characterized according to RT (retention time), accurate mass of primary mass spectrometry and secondary mass spectrometry, with  $\Delta\text{RT} \leq 0.5$  min and  $\Delta\text{m/s} \leq 5$  ppm. Substances that were not in the local library were annotated by accurate mass alignment with primary and secondary mass spectra of substances in public databases AraCyc, BioCyc, Human Metabolome Database, KEGG, Lipid Maps, and PlantCyc with  $\Delta\text{m/s} \leq 5$  ppm. Solvent mass spectra were used to remove background noise. QC was used for peak alignment and correction.

## Data analysis

Metabolite taxonomic statistics were performed using Microsoft Office Excel 2010. The Pearson correlation coefficient between QCs was calculated using the gpubr package in R3.6.2 and visualized with the ggplot2 package. Cluster analysis was performed on multiple samples according to the composition and content of metabolites. The fpc and vegan packages were used to determine the number of clusters according to K-means. The optimal k value was determined based on Calinski criterion. The distance matrix between samples was calculated via the cluster package based on ward.D2 in the pedigree clustering. Ape package was used for visualization. DESeq2 was used to calculate log<sub>2</sub> (Fold Change) and p value between groups after variance analysis. The p value was corrected via Benjamini-hochberg to obtain adj-p. DAMs (differentially accumulated metabolites) were screened as Fold change > 2 and adj-p < 0.05. Volcano plots were created using the ggplot2 package.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (32072631), the Chinese Academy of Agricultural Sciences through the Agricultural Science and Technology Innovation Program (CAASTIP-2017-TRICAAS), China Agriculture Research System of MOF and MARA (CARS-19) to LC.

## Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary Information** accompanies this paper at (<http://www.maxapress.com/article/doi/10.48130/BPR-2022-0009>)

## Dates

Received 3 April 2022; Accepted 6 May 2022; Published online 1 June 2022

## REFERENCES

- Dai W, Qi D, Yang T, Lv H, Guo L, et al. 2015. 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.). *Journal of Agricultural and Food Chemistry* 63:9869–78
- Zhang L, Tai Y, Wang Y, Meng Q, Yang Y, et al. 2017. The proposed biosynthesis of procyanidins by the comparative chemical analysis of five *Camellia* species using LC-MS. *Scientific Reports* 7:46131
- Chen L, Yu F, Tong Q. 2000. Discussion on phylogenetic classification and evolution of Sect. *Thea*. *Journal of Tea Sciences* 20:89–94
- Yu F. 2016. Ancient Tea Plants in China. Kunming: Yunnan Science and Technology Press. pp. 33
- Liu Y, Zhao G, Li X, Shen Q, Wu Q, et al. 2020. Comparative analysis of phenolic compound metabolism among tea plants in the section *Thea* of the genus *Camellia*. *Food Research International* 135:109276
- Teng J, Yan C, Zeng W, Zhang Y, Zeng Z, et al. 2020. Purification and characterization of theobromine synthase in a theobromine-enriched wild tea plant (*Camellia gymnogyna* Chang) from Dayao Mountain, China. *Food Chemistry* 311:125875
- Wildermuth MC. 2006. Variations on a theme: synthesis and modification of plant benzoic acids. *Current Opinion in Plant Biology* 9:288–96
- Liu Z, Bruins ME, de Bruijn WJC, Vincken JP. 2020. A comparison of the phenolic composition of old and young tea leaves reveals a decrease in flavanols and phenolic acids and an increase in flavonols upon tea leaf maturation. *Journal of Food Composition and Analysis* 86:103385
- Fan K, Zhang Q, Liu M, Ma L, Shi Y, et al. 2019. Metabolomic and transcriptional analyses reveal the mechanism of C, N allocation from source leaf to flower in tea plant (*Camellia sinensis* L.). *Journal of Plant Physiology* 232:200–8
- Jiang H, Yu F, Qin L, Zhang N, Cao Q, et al. 2019. Dynamic change in amino acids, catechins, alkaloids, and gallic acid in six types of tea processed from the same batch of fresh tea (*Camellia sinensis* L.) leaves. *Journal of Food Composition and Analysis* 77:28–38
- Meinhart AD, Caldeirão L, Damin FM, Filho JT, Godoy HT. 2018. Analysis of chlorogenic acids isomers and caffeic acid in 89 herbal infusions (tea). *Journal of Food Composition and Analysis* 73:76–82

12. Lee VSY, Dou J, Chen RJY, Lin RS, Lee MR, et al. 2008. Massive accumulation of gallic acid and unique occurrence of myricetin, quercetin, and kaempferol in preparing old oolong tea. *Journal of Agricultural and Food Chemistry* 56:7950
13. Meng X, Li N, Zhu H, Wang D, Yang C, et al. 2019. Plant resources, chemical constituents, and bioactivities of tea plants from the genus *Camellia* section *Thea*. *Journal of Agricultural and Food Chemistry* 67:5318–49
14. Hao W, Ma J, Ma C, Jin J, Chen L. 2019. The complete chloroplast genome sequence of *Camellia tachangensis* F. C. Zhang (Theaceae). *Mitochondrial DNA Part B: Resources* 4:3344–45
15. Gao D, Zhang Y, Yang C, Chen K, Jiang H. 2009. Phenolic antioxidants from green tea produced from *Camellia crassicolumna* var. *multiplex*. *Journal of Agricultural and Food Chemistry* 57:586–90
16. Zhu B, Chen L, Lu M, Zhang J, Han J, et al. 2019. Caffeine content and related gene expression: novel insight into caffeine metabolism in *Camellia* plants containing low, normal, and high caffeine concentrations. *Journal of Agricultural and Food Chemistry* 67:3400–11
17. Yang C, Zhang Y, Gao D, Chen K, Jiang H. 2008. The evaluation of *Camellia taliensis* resources and the origin of cultivated large-leave tea plant. *Tea science and Technology* 3:1–4
18. Zhang H. 1984. A revision on the tea resource plants. *Acta Scientiarum Naturalium Universitatis Sunyatseni* 23:3–14
19. Gao D, Zhang Y, Yang C, Chen K, Jiang H. 2008. Phenolic Antioxidants from green tea produced from *Camellia taliensis*. *Journal of Agricultural and Food Chemistry* 56:7517–21
20. Yan Z, Huang J, Wang D. 2020. Research progress in natural theobromine-enriched tea plants. *Guangdong Agricultural Sciences* 47:37–44
21. Zhu W, Xu J, Ge Y, Cao H, Ge X, et al. 2014. Epigallocatechin-3-gallate (EGCG) protects skin cells from ionizing radiation via heme oxygenase-1 (HO-1) overexpression. *Journal of Radiation Research* 55:1056–65
22. Cai J, Jozwiak A, Holoidovsky L, Meijler MM, Meir S, et al. 2021. Glycosylation of *N*-hydroxy-pipecolic acid equilibrates between systemic acquired resistance response and plant growth. *Molecular Plant* 14:440–55
23. Li D, Wang J, Liu Y, Li Y, Zhang Z. 2021. Expanded analyses of the functional correlations within structural classifications of glycoside hydrolases. *Computational and Structural Biotechnology Journal* 19:5931–42
24. Cao S, Zhang M, Yuan M, Yang D, Zhao M, et al. 2022. The pharmaceutical excipient PEG400 affect the absorption of baicalein in Caco-2 monolayer model by interacting with UDP-glucuronosyltransferases and efflux transport proteins. *Pharmacology research & perspectives* 10:e00928
25. Zhu J, Xu Q, Zhao S, Xia X, Yan X, et al. 2020. Comprehensive co-expression analysis provides novel insights into temporal variation of flavonoids in fresh leaves of the tea plant (*Camellia sinensis*). *Plant Science* 290:110306
26. Wu C, Xu H, Héritier J, Andlauer W. 2012. Determination of catechins and flavonol glycosides in Chinese tea varieties. *Food Chemistry* 32:144–49
27. Dai W, Tan J, Lu M, Xie D, Li P, et al. 2016. Nontargeted modification-specific metabolomics investigation of glycosylated secondary metabolites in tea (*Camellia sinensis* L.) based on liquid chromatography-high resolution mass spectrometry. *Journal of Agricultural and Food Chemistry* 64:6783–90
28. Li X, Liu G, Zhang W, Zhou YL, Ling T, et al. 2018. Novel flavoalkaloids from white tea with inhibitory activity against the formation of advanced glycation end products. *Journal of Agricultural and Food Chemistry* 66:4621–29



Copyright: © 2022 by the author(s). Exclusive Licensee Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.